

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Murphy, Brian R.
Collins, Peter L.
Durbin, Anna P.
Skiadopoulos, Mario H.
Tao, Tao

(ii) TITLE OF INVENTION: PRODUCTION OF ATTENUATED PARAINFLUENZA VIRUS VACCINES FROM CLONED NUCLEOTIDE SEQUENCE

(iii) NUMBER OF SEQUENCES: 74

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
(B) STREET: Two Embarcadero Center, 8th Floor
(C) CITY: San Francisco
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 94111

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 09/
(B) FILING DATE: 22-MAY-1998
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 60/059,385
(B) FILING DATE: 19-SEP-1997

(viii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 60/047,575
(B) FILING DATE: 23-MAY-1997

(ix) ATTORNEY/AGENT INFORMATION:

(A) NAME: King, Jeffrey J.
(B) REGISTRATION NUMBER: 38,515
(C) REFERENCE/DOCKET NUMBER: 17634-000320

(x) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 206-467-9600
(B) TELEFAX: 415-576-0300

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15669 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCTGGGTACCGGGCCCGTCGACGCGTATATAGTTCCCTCCTTTCAGCAAAA AACCCCTCAA

GACCCGTTA GAGGCCCAA GGGTTATGC TACTGCAGGC TCTCCCTTAG CCATCCGAGT	120
GGACGTGCGT CCTCCTTCGG ATGCCAGGT CGGACCGCGA GGAGGTGGAG ATGCCATGCC	180
GACCCACCAA ACAAGAGAAAG AAACCTGTCT GGGAAATATAA ATTTAACCTT AAATTAACCTT	240
AGGATTAAAG ACATTGACTA GAAGGTCAAG AAAAGGGAAC TCTATAATTT CAAAAATGTT	300
GAGCCTATTT GATACATTAA ATGCACGTAG GCAAGAAAAC ATAACAAAAT CAGCCGGTGG	360
AGCTATCATT CCTGGACAGA AAAATACTGT CTCTATATTC GCCCTGGAC CGACAATAAC	420
TGATGATAAT GAGAAAATGA CATTAGCTCT TCTATTTCTA TCTCATTACAC TAGATAATGA	480
GAAACAACAT GCACAAAGGG CAGGGTTCTT GGTGTCTTA TTGTCAATGG CTTATGCCAA	540
TCCAGAGCTC TACCTAACAA CAAATGGAAG TAATGCAGAT GTCAAGTATG TCATATACAT	600
GATTGAGAAA GATCTAAAAC GGCAAAAGTA TGGAGGATTT GTGGTTAAGA CGAGAGAGAT	660
GATATATGAA AAGACAACTG ATTGGATATT TGGAAGTGAC CTGGATTATG ATCAGGAAAC	720
TATGTTGCAG AACGGCAGGA ACAATTCAAC AATTGAAGAC CTTGTCCACA CATTGGGTA	780
TCCATCATGT TTAGGAGCTC TTATAATACA GATCTGGATA GTTCTGGTCA AAGCTATCAC	840
TAGTATCTCA GGGTTAAGAA AAGGCTTTT CACCGATTG GAAGCTTCA GACAAGATGG	900
AACAGTGCAG GCAGGGCTGG TATTGAGCGG TGACACAGTG GATCAGATTG GGTCAATCAT	960
GCGGTCTCAA CAGAGCTTGG TAACTCTTAT GGTTGAAACA TTAATAACAA TGAATACCAG	1020
CAGAAATGAC CTCACAAACCA TAGAAAAGAA TATACAAATT GTGGCAACT ACATAAGAGA	1080
TGCAGGTCTC GCTTCATTCT TCAATACAAT CAGATATGGA ATTGAGACCA GAATGGCAGC	1140
TTTGACTCTA TCCACTCTCA GACCAGATAT CAATAGATTA AAAGCTTGA TGGAACGTGA	1200
TTTATCAAAG GGACCACGCG CTCCTTCAT CTGTATCCTC AGAGATCCTA TACATGGTGA	1260
GTTCGCACCA GGCAACTATC CTGCCATATG GAGCTATGCA ATGGGGTGG CAGTTGTACA	1320
AAATAGAGCC ATGCAACAGT ATGTGACGGG AAGATCATAT CTAGACATTG ATATGTTCCA	1380
GCTAGGACAA GCAGTAGCAC GTGATGCCGA AGCTCAAATG AGCTAACAC TGGAAGATGA	1440
ACTTGGAGTG ACACACGAAT CTAAGAAAG CTTGAAGAGA CATATAAGGA ACATAAACAG	1500
TTCAGAGACA TCTTCCACA AACCGACAGG TGGATCAGCC ATAGAGATGG CAATAGATGA	1560
AGAGCCAGAA CAATTGAAAC ATAGAGCAGA TCAAGAACAA AATGGAGAAC CTCAACATC	1620
CATAATTCAA TATGCCTGGG CAGAAGAAA TAGAAGCGAT GATCAGACTG AGCAAGCTAC	1680
AGAATCTGAC AATATCAAGA CCGAACAAACA AAACATCAGA GACAGACTAA ACAAGAGACT	1740
CAACGACAAG AAGAAACAAA GCAGTCAACC ACCCACTAAT CCCACAAACA GAACAAACCA	1800
GGACGAAATA GATGATCTGT TTAACGCATT TGGAAGCAAC TAATCGAAC AACATTTAA	1860
TCTAAATCAA TAATAAATAA GAAAAACTTA GGATTAAAGA ATCCTATCAT ACCGGAATAT	1920
AGGGTGGTAA ATTTAGAGTC TGCTTGAAAC TCAATCAATA GAGAGTTGAT GGAAAGCGAT	1980
GCTAAAAACT ATCAAATCAT GGATTCTGG GAAGAGGAAT CAAGAGATAA ATCAAACATAAT	2040
ATCTCCTCGG CCCTCAACAT CATTGAATTC ATACTCAGCA CCGACCCCCA AGAAGACTTA	2100
TCGGAAAACG ACACAATCAA CACAAGAACC CAGCAACTCA GTGCCACCAT CTGTCAACCA	2160

GAAATCAAAC	CAACAGAAC	AAAGTGAGAAA	GATAGTGGAT	CAACTGACAA	AAATAGACAG	2220
TCTGGGTCA	CACACGAATG	TACAACAGAA	GCAAAAGATA	GAAATATTGA	TCAGGAAACT	2280
GTACAGAGAG	GACCTGGGAG	AAGAAGCAGC	TCAGATAGTA	GAGCTGAGAC	TGTGGTCTCT	2340
GGAGGAATCC	CCAGAAGCAT	CACAGATTCT	AAAAATGGAA	CCCAAAACAC	GGAGGATATT	2400
GATCTCAATG	AAATTAGAAA	GATGGATAAG	GAECTCTATTG	AGGGGAAAAT	GCGACAATCT	2460
GCAAATGTT	CAAGCGAGAT	ATCAGGAAGT	GATGACATAT	TTACAACAGA	ACAAAGTAGA	2520
AACAGTGATC	ATGGAAGAAG	CCTGGAATCT	ATCAGTACAC	CTGATACAAG	ATCAATAAGT	2580
GTTGTTACTG	CTGCAACACC	AGATGATGAA	GAAGAAATAC	TAATGAAAAA	TAGTAGGACA	2640
AAGAAAAGTT	CTTCAACACA	TCAAGAAGAT	GACAAAAGAA	TTAAAAAAGG	GGGAAAAGGG	2700
AAAGACTGGT	TTAAGAAATC	AAAAGATACC	GACAACCAGA	TACCAACATC	AGACTACAGA	2760
TCCACATCAA	AAGGGCAGAA	GAAAATCTCA	AAGACAACAA	CCACCAACAC	CGACACAAAG	2820
GGGCAAACAG	AAATACAGAC	AGAATCATCA	GAAACACAAT	CCTCATCATG	GAATCTCATC	2880
ATCGACAACA	ACACCGACCG	GAACGAACAG	ACAAGCACAA	CTCCTCCAAC	AAACAACCTCC	2940
AGATCAACTT	ATACAAAAGA	ATCGATCCGA	ACAAACTCTG	AATCCAAACC	CAAGACACAA	3000
AAGACAAATG	GAAAGGAAAG	GAAGGATACA	GAAGAGAGCA	ATCGATTTAC	AGAGAGGGCA	3060
ATTACTCTAT	TGCAGAATCT	TGGTGTAATT	CAATCCACAT	CAAAACTAGA	TTTATATCAA	3120
GACAAACGAG	TTGTATGTGT	AGCAAATGTA	CTAAACAATG	TAGATACTGC	ATCAAAGATA	3180
GATTCCTGG	CAGGATTAGT	CATAGGGTT	TCAATGGACA	ACGACACAAA	ATTAACACAG	3240
ATACAAAATG	AAATGCTAA	CCTCAAAGCA	GATCTAAAGA	AAATGGACGA	ATCACATAGA	3300
AGATTGATAG	AAAATCAAAG	AGAACAACTG	TCATTGATCA	CGTCACTAAT	TTCAAATCTC	3360
AAAATTATGA	CTGAGAGAGG	AGGAAAGAAA	GACCAAAATG	AATCCAATGA	GAGAGTATCC	3420
ATGATCAAAA	CAAATTGAA	AGAAGAAAAG	ATCAAGAAGA	CCAGGTTGA	CCCACCTTATG	3480
GAGGCACAAG	GCATTGACAA	GAATATACCC	GATCTATATC	GACATGCAGG	AGATACACTA	3540
GAGAACGGATG	TACAAGTTAA	ATCAGAGATA	TTAAGTTCAT	ACAATGAGTC	AAATGCAACA	3600
AGACTAATAC	CCAAAAAAAGT	GAGCAGTACA	ATGAGATCAC	TAGTTGCAGT	CATCAACAAAC	3660
AGCAATCTCT	CACAAAGCAC	AAAACAATCA	TACATAAACG	AACTCAAACG	TTGCAAAAT	3720
GATGAAGAAG	TATCTGAATT	AATGGACATG	TTCAATGAAG	ATGTCAACAA	TTGCCAATGA	3780
TCCAACAAAG	AAACGACACC	GAACAAACAG	ACAAGAAACA	ACAGTAGATC	AAAACCTGTC	3840
AACACACACA	AAATCAAGCA	GAATGAAACA	ACAGATATCA	ATCAATATAC	AAATAAGAAA	3900
AACTTAGGAT	TAAAGAATAA	ATTAATCCTT	GTCCAAAATG	AGTATAACTA	ACTCTGCAAT	3960
ATACACATTC	CCAGAATCAT	CATTCTCTGA	AAATGGTCAT	ATAGAACCAT	TACCACTCAA	4020
AGTCAATGAA	CAGAGGAAAG	CAGTACCCCA	CATTAGAGTT	GCCAAGATCG	GAAATCCACC	4080
AAAACACGGA	TCCCGGTATT	TAGATGTCTT	CTTACTCGGC	TTCTTCGAGA	TGGAACGAAT	4140
CAAAGACAAA	TACGGGAGTG	TGAATGATCT	CGACAGTGAC	CCGAGTTACA	AAGTTGTGG	4200
CTCTGGATCA	TTACCAATCG	GATTGGCTAA	GTACACTGGG	AATGACCAGG	AATTGTTACA	4260

AGCCGCAACC	AAACTGGATA	TAGAAGTGAG	AAGAACAGTC	AAAGCGAAAG	AGATGGTTGT	4320
TTACACGGTA	CAAAATATAA	AACCAGAACT	GTACCCATGG	TCCAATAGAC	TAAGAAAAGG	4380
AATGCTGTTG	GATGCCAACA	AAGTTGCTCT	TGCTCCTCAA	TGTCTTCCAC	TAGATAGGAG	4440
CATAAAATTG	AGAGTAATCT	TCGTGAATTG	TACGGCAATT	GGATCAATAA	CCTTGTCAA	4500
AATTCCCTAAG	TCAATGGCAT	CACTATCTCT	ACCCAACACA	ATATCAATCA	ATCTGCAGGT	4560
ACACATAAAA	ACAGGGGTTTC	AGACTGATTC	TAAAGGGATA	GTTCAAATTG	TGGATGAGAA	4620
AGGCAGAAAAA	TCACTGAATT	TCATGGTCCA	TCTCGGATTG	ATCAAAAGAA	AAGTAGGCAG	4680
AATGTACTCT	GTTGAATACT	GTAAACAGAA	AATCGAGAAA	ATGAGATTGA	TATTTTCTTT	4740
AGGACTAGTT	GGAGGAATCA	GTCTTCATGT	CAATGCAACT	GGGTCATAT	CAAAAACACT	4800
AGCAAGTCAG	CTGGTATTCA	AAAGAGAGAT	TTGTTATCCT	TTAATGGATC	TAAATCCGCA	4860
TCTCAATCTA	GTTATCTGGG	CTTCATCAGT	AGAGATTACA	AGAGTGGATG	CAATTTCCA	4920
ACCTTCTTTA	CCTGGCGAGT	TCAGATACTA	TCCTAATATT	ATTGCAAAAG	GAGTTGGAA	4980
AATCAAACAA	TGGAACTAGT	AATCTCTATT	TTAGTCCGGA	CGTATCTATT	AAGCCGAAGC	5040
AAATAAAGGA	TAATCAAAAAA	CTTAGGACAA	AAGAGGTCAA	TACCAACAAAC	TATTAGCAGT	5100
CACACTCGCA	AGAATAAGAG	AGAAGGGACC	AAAAAAGTCA	AATAGGAGAA	ATCAAAACAA	5160
AAGGTACAGA	ACACCAGAAC	AACAAAATCA	AAACATCCAA	CTCACTCAAA	ACAAAAATTCA	5220
CAAAAGAGAC	CGGCAACACA	ACAAGCACTG	AACACAATGC	CAACTTCAT	ACTGCTAATT	5280
ATTACAACCA	TGATCATGGC	ATCTTCTGC	CAAATAGATA	TCACAAAACT	ACAGCACGTA	5340
GGTGTATTGG	TCAACAGTCC	CAAAGGGATG	AAGATATCAC	AAAACTTGA	AACAAGATAT	5400
CTAATTGGA	GCCTCATACC	AAAAATAGAA	GACTCTAACT	CTTGTGGTGA	CCAACAGATC	5460
AAGCAATACA	AGAAGTTATT	GGATAGACTG	ATCATCCCTT	TATATGATGG	ATTAAGATTA	5520
CAGAAAGATG	TGATAGTAAC	CAATCAAGAA	TCCAATGAAA	ACACTGATCC	CAGAACAAAA	5580
CGATTCTTTG	GAGGGGTAAT	TGGAACCATT	GCTCTGGGAG	TAGCAACCTC	AGCACAAATT	5640
ACAGCGGCAG	TTGCTCTGGT	TGAAGCCAAG	CAGGCAAGAT	CAGACATCGA	AAAACCTCAA	5700
GAAGCAATTAA	GGGACACAAA	CAAAGCAGTG	CAGTCAGTTC	AGAGCTCCAT	AGGAAATTAA	5760
ATAGTAGCAA	TTAAATCAGT	CCAGGATTAT	GTAAACAAAG	AAATCGTGCC	ATCGATTGCG	5820
AGGCTAGGTT	GTGAAGCAGC	AGGACTCAA	TTAGGAATTG	CATTAACACA	GCATTACTCA	5880
GAATTAACAA	ACATATTG	TGATAACATA	GGATCGTTAC	AAGAAAAGG	AATAAAATTAA	5940
CAAGGTATAG	CATCATTATA	CCGCACAAAT	ATCACAGAAA	TATTCACAAAC	ATCAACAGTT	6000
GATAAAATATG	ATATCTATGA	TCTGTTATTT	ACAGAATCAA	TAAAGGTGAG	AGTTATAGAT	6060
GTTGACTTGA	ATGATTACTC	AATCACCCCTC	CAAGTCAGAC	TCCCTTATT	AACTAGGCTG	6120
CTGAACACTC	AGATCTACAA	AGTAGATTCC	ATATCATATA	ACATCCAAA	CAGAGAATGG	6180
TATATCCCTC	TTCCCAGCCA	TATCATGACG	AAAGGGCAT	TTCTAGGTGG	AGCAGACGTC	6240
AAAGAATGTA	TAGAAGCATT	CAGCAGCTAT	ATATGCCCTT	CTGATCCAGG	ATTTGTATTA	6300
AACCATGAAA	TAGAGAGCTG	CTTATCAGGA	AACATATCCC	AATGTCCAAG	AACAACGGTC	6360

ACATCAGACA TTGTTCCAAG ATATGCATTT GTCAATGGAG GAGTGGTTGC AAACTGTATA	6420
ACAACCACCT GTACATGCAA CGGAATTGGT AATAGAACATCA ATCAACCACC TGATCAAGGA	6480
GTAAAAATTA TAACACATAA AGAATGTAGT ACAATAGGTA TCAACGGAAT GCTGTTCAAT	6540
ACAAATAAAG AAGGAACCTCT TGCATTCTAT ACACCAAATG ATATAACACT AAACAATTCT	6600
GTTGCACTTG ATCCAATTGA CATATCAATC GAGCTCAACA AGGCCAAATC AGATCTAGAA	6660
GAATCAAAAG AATGGATAAG AAGGTCAAAT CAAAAACTAG ATTCTATTGG AAATTGGCAT	6720
CAATCTAGCA CTACAATCAT AATTATTTG ATAATGATCA TTATATTGTT TATAATTAAAT	6780
ATAACGATAA TTACAATTGC AATTAAGTAT TACAGAATTC AAAAGAGAAA TCGAGTGGAT	6840
CAAAATGACA AGCCATATGT ACTAACAAAC AAATAACATA TCTACAGATC ATTAGATATT	6900
AAAATTATAA AAAACTTAGG AGTAAAGTTA CGCAATCCAA CTCTACTCAT ATAATTGAGG	6960
AAGGACCCAA TAGACAAATC CAAATTCGAG ATGGAATACT GGAAGCATAAC CAATCACGGA	7020
AAGGATGCTG GTAATGAGCT GGAGACGTCT ATGGCTACTC ATGGCAACAA GCTCACTAAT	7080
AAGATAATAT ACATATTATG GACAATAATC CTGGTGTAT TATCAATAGT CTTCATCATA	7140
GTGCTAATTA ATTCCATCAA AAGTGAAAAG GCCCACGAAT CATTGCTGCA AGACATAAAAT	7200
AATGAGTTA TGGAATTAC AGAAAAGATC CAAATGGCAT CGGATAATAC CAATGATCTA	7260
ATACAGTCAG GAGTGAATAC AAGGCTTCTT ACAATTCAAG GTCATGTCCA GAATTACATA	7320
CCAATATCAT TGACACAACA GATGTCAGAT CTTAGGAAAT TCATTAGTGA AATTACAATT	7380
AGAAAATGATA ATCAAGAAGT GCTGCCACAA AGAATAACAC ATGATGTAGG TATAAAACCT	7440
TTAAATCCAG ATGATTTTG GAGATGCACG TCTGGTCTTC CATCTTAAT GAAAACCTCA	7500
AAAATAAGGT TAATGCCAGG GCCGGGATTA TTAGCTATGC CAACGACTGT TGATGGCTGT	7560
GTTAGAACTC CGTCTTTAGT TATAATGAT CTGATTTATG CTTATACCTC AAATCTAATT	7620
ACTCGAGGTT GTCAGGATAT AGGAAAATCA TATCAAGTCT TACAGATAGG GATAATAACT	7680
GTAAAATCAG ACTTGGTACC TGACTTAAAT CCTAGGATCT CTCATACCTT TAACATAAAAT	7740
GACAATAGGA AGTCATGTT TCTAGCACTC CTAATATAG ATGTATATCA ACTGTGTTCA	7800
ACTCCAAAG TTGATGAAAG ATCAGATTAT GCATCATCAG GCATAGAAGA TATTGTACTT	7860
GATATTGTCA ATTATGATGG TTCAATCTCA ACAACAAGAT TTAAGAATAA TAACATAAGC	7920
TTTGATCAAC CATATGCTGC ACTATACCA TCTGTTGGAC CAGGGATATA CTACAAAGGC	7980
AAAATAATAT TTCTCGGGTA TGGAGGTCTT GAACATCCAA TAAATGAGAA TGTAATCTGC	8040
AACACAACAG GGTGCCCGG GAAAACACAG AGAGACTGTA ATCAAGCATE TCATAGTACT	8100
TGGTTTCAG ATAGGAGGAT GGTCAACTCC ATCATTGTTG TTGACAAAGG CTTAAACTCA	8160
ATTCCAAAAT TGAAAGTATG GACGATATCT ATGCGACAAA ATTACTGGGG GTCAGAAGGA	8220
AGGTTACTTC TACTAGGTAA CAAGATCTAT ATATATACAA GATCTACAAG TTGGCATAGC	8280
AAGTTACAAT TAGGAATAAT TGATATTACT GATTACAGTG ATATAAGGAT AAAATGGACA	8340
TGGCATAATG TGCTATCAAG ACCAGGAAAC AATGAATGTC CATGGGGACA TTCATGTCCA	8400
GATGGATGTA TAACAGGAGT ATATACTGAT GCATATCCAC TCAATCCCAC AGGGAGCATT	8460

GTGTCATCTG TCATATTAGA CTCACAAAAA TCGAGAGTGA ACCCAGTCAT AACTTACTCA	8520
ACAGCAACCG AAAGAGTAAA CGAGCTGGCC ATCCTAAACA GAACACTCTC AGCTGGATAT	8580
ACAACAACAA GCTGCATTAC ACACTATAAC AAAGGATATT GTTTCATAT AGTAGAAATA	8640
AATCATAAAA GCTTAAACAC ATTTCAACCC ATGTTGTTCA AAACAGAGAT TCCAAAAGC	8700
TGCAGTTAAT CATAATTAAC CATAATATGC ATCAATCTAT CTATAATACA AGTATATGAT	8760
AAGTAATCAG CAATCAGACA ATAGACAAAA GGGAAATATA AAAAACCTAG GAGCAAAGCG	8820
TGCTCGGAA ATGGACACTG AATCTAACAA TGGCACTGTA TCTGACATAC TCTATCCTGA	8880
GTGTCACCTT AACTCTCCTA TCGTAAAGG TAAAATAGCA CAATTACACA CTATTATGAG	8940
TCTACCTCAG CCTTATGATA TGGATGACGA CTCAATACTA GTTATCACTA GACAGAAAAT	9000
AAAACCTTAAT AAATTGGATA AAAGACAACG ATCTATTAGA AGATTAAAAT TAATATTAAC	9060
TGAAAAAGTG AATGACTTAG GAAAATACAC ATTTATCAGA TATCCAGAAA TGTCAAAAGA	9120
AATGTTCAAA TTATATATAC CTGGTATTAA CAGTAAAGTG ACTGAATTAT TACTTAAAGC	9180
AGATAGAACAA TATAGTCAAA TGACTGATGG ATTAAGAGAT CTATGGATTA ATGTGCTATC	9240
AAAATTAGCC TCAAAAAATG ATGGAAGCAA TTATGATCTT AATGAAGAAA TTAATAATAT	9300
ATCGAAAGTT CACACAACCT ATAAATCAGA TAAATGGTAT AATCCATTCA AAACATGGTT	9360
TACTATCAAG TATGATATGA GAAGATTACA AAAAGCTCGA AATGAGATCA CTTTTAATGT	9420
TGGGAAGGAT TATAACTTGT TAGAAGACCA GAAGAATTTC TTATTGATAC ATCCAGAATT	9480
GGTTTTGATA TTAGATAAAC AAAACTATAA TGGTTATCTA ATTACTCCTG AATTAGTATT	9540
GATGTATTGT GACGTAGTCG AAGGCCGATG GAATATAAGT GCATGTGCTA AGTTAGATCC	9600
AAAATTACAA TCTATGTATC AGAAAGGTA TAACCTGTGG GAAGTGATAG ATAAATTGTT	9660
TCCAATTATG GGAGAAAAGA CATTGATGT GATATCGTTA TTAGAACACC TTGCATTATC	9720
CTTAATTCAA ACTCATGATC CTGTTAAACA ACTAAGAGGA GCTTTTTAA ATCATGTGTT	9780
ATCCGAGATG GAATTAATAT TTGAATCTAG AGAATCGATT AAGGAATTTC TGAGTGTAGA	9840
TTACATTGAT AAAATTAG ATATATTAA TAAGTCTACA ATAGATGAAA TAGCAGAGAT	9900
TTTCTCTTT TTTAGAACAT TTGGGCATCC TCCATTAGAA GCTAGTATTG CAGCAGAAAA	9960
GGTTAGAAAA TATATGTATA TTGGAAAACA ATTAAAATTG GACACTATTA ATAAATGTCA	10020
TGCTATCTTC TGTACAATAA TAATTAACGG ATATAGAGAG AGGCATGGTG GACAGTGGCC	10080
TCCTGTGACA TTACCTGATC ATGCACACGA ATTCACTCATA AATGCTTACG GTTCAAACTC	10140
TGCGATATCA TATGAAAATG CTGTTGATTA TTACCAAGAGC TTTATAGGAA TAAAATTCAA	10200
TAAATTCTATA GAGCCTCAGT TAGATGAGGA TTTGACAATT TATATGAAAG ATAAAGCATT	10260
ATCTCCAAAA AAATCAAATT GGGACACAGT TTATCCTGCA TCTAATTTAC TGTACCGTAC	10320
TAACGCATCC AACGAATCAC GAAGATTAGT TGAAGTATTG ATAGCAGATA GTAAATTGAA	10380
TCCTCATCAG ATATTGGATT ATGTAGAATC TGGGGACTGG TTAGATGATC CAGAATTAA	10440
TATTTCTTAT AGTCTTAAAG AAAAGAGAT CAAACAGGAA GGTAGACTCT TTGCAAAAT	10500
GACATACAAA ATGAGAGCTA CACAAGTTT ATCAGAGACC CTACTTGCAA ATAACATAGG	10560

AAAATTCTTT	CAAGAAAATG	GGATGGTGAA	GGGAGAGATT	GAATTACTTA	AGAGATTAAC	10620
AACCATAATCA	ATATCAGGAG	TTCCACGGTA	TAATGAAGTG	TACAATAATT	CTAAAAGCCA	10680
TACAGATGAC	CTTAAAACCT	ACAATAAAAT	AAGTAATCTT	AATTTGTCTT	CTAATCAGAA	10740
ATCAAAGAAA	TTTGAATTCA	AGTCAACGGA	TATCTACAAT	GATGGATACG	AGACTGTGAG	10800
CTGTTTCCTA	ACAACAGATC	TCAAAAAATA	CTGTCTTAAT	TGGAGATATG	AATCAACAGC	10860
TCTATTGGA	GAAACTTGCA	ACCAAATATT	TGGATTAAAT	AAATTGTTTA	ATTGGTTACA	10920
CCCTCGTCTT	GAAGGAAGTA	CAATCTATGT	AGGTGATCCT	TACTGTCCTC	CATCAGATAA	10980
AGAACATATA	TCATTAGAGG	ATCACCCCTGA	TTCTGGTTTT	TACGTTCATATA	ACCCAAGAGG	11040
GGGTATAGAA	GGATTTGTC	AAAAATTATG	GACACTCATA	TCTATAAGTG	CAATACATCT	11100
AGCAGCTGTT	AGAATAGGCG	TGAGGGTGAC	TGCAATGGTT	CAAGGAGACA	ATCAAGCTAT	11160
AGCTGTAACC	ACAAGAGTAC	CCAACAATTA	TGACTACAGA	GTAAAGAAGG	AGATAGTTA	11220
TAAAGATGTA	GTGAGATTTT	TTGATTCTATT	AAGAGAAGTG	ATGGATGATC	TAGGTCTGATGA	11280
ACTTAAATTA	AATGAAACGA	TTATAAGTAG	CAAGATGTT	ATATATAGCA	AAAGAATCTA	11340
TTATGATGGG	AGAATTCTTC	CTCAAGCTCT	AAAAGCATTA	TCTAGATGTG	TCTTCTGGTC	11400
AGAGACAGTA	ATAGACGAAA	CAAGATCAGC	ATCTTCAAAT	TTGGCAACAT	CATTTGCCAAA	11460
AGCAATTGAG	AATGGTTATT	CACCTGTTCT	AGGATATGCA	TGCTCAATT	TTAAGAATAT	11520
TCAACAACTA	TATATTGCC	TTGGGATGAA	TATCAATCCA	ACTATAACAC	AGAATATCAG	11580
AGATCAGTAT	TTTAGGAATC	CAAATTGGAT	GCAATATGCC	TCTTTAATAC	CTGCTAGTGT	11640
TGGGGGATTC	AATTACATGG	CCATGTCAAG	ATGTTTGTA	AGGAATATTG	GTGATCCATC	11700
AGTTGCCGCA	TTGGCTGATA	TTAAAAGATT	TATTAAGGCG	AATCTATTAG	ACCGAAGTGT	11760
TCTTTATAGG	ATTATGAATC	AAGAACCAAG	TGAGTCATCT	TTTTTGGACT	GGGCTTCAGA	11820
TCCATATTCA	TGCAATTTCAC	CACAATCTCA	AAATATAACC	ACCATGATAA	AAAATATAAC	11880
AGCAAGGAAT	GTATTACAAG	ATTCACCAAA	TCCATTATTA	TCTGGATTAT	TCACAAATAC	11940
AATGATAGAA	GAAGATGAAG	AATTAGCTGA	GTTCTGATG	GACAGGAAGG	TAATTCTCCC	12000
TAGAGTTGCA	CATGATATTC	TAGATAATTC	TCTCACAGGA	ATTAGAAATG	CCATAGCTGG	12060
AATGTTAGAT	ACGACAAAAT	CACTAATTG	GGTTGGCATA	AATAGAGGAG	GAUTGACATA	12120
TAGTTGTTG	AGGAAAATCA	GTAATTACGA	TCTAGTACAA	TATGAAACAC	TAAGTAGGAC	12180
TTTGCACTA	ATTGTAAGTG	ATAAAATCAA	GTATGAAGAT	ATGTGTTCGG	TAGACCTTGC	12240
CATAGCATTG	CGACAAAAGA	TGTGGATTCA	TTTATCAGGA	GGAAGGATGA	TAAGTGGACT	12300
TGAAACGCC	GACCCATTAG	AATTACTATC	TGGGGTAGTA	ATAACAGGAT	CAGAACATTG	12360
TAAAATATGT	TATTCTTCAG	ATGGCACAAA	CCCATATACT	TGGATGTATT	TACCCGGTAA	12420
TATCAAAATA	GGATCAGCAG	AAACAGGTAT	ATCGTCATTA	AGAGTTCCCTT	ATTTTGGATC	12480
AGTCACTGAT	GAAAGATCTG	AAGCACAATT	AGGATATATC	AAGAATCTTA	GTAAACCTGC	12540
AAAAGCCGCA	ATAAGAATAG	CAATGATATA	TACATGGCA	TTTGGTAATG	ATGAGATATC	12600
TTGGATGGAA	GCCTCACAGA	TAGCACAAAC	ACGTGCAAAT	TTTACACTAG	ATAGTCTCAA	12660

AATTTTAACA	CCGGTAGCTA	CATCAACAAA	TTTATCACAC	AGATTAAAGG	ATACTGCAAC	12720
TCAGATGAAA	TTCTCCAGTA	CATCATTGAT	CAGAGTCAGC	AGATTCTAA	CAATGTCCAA	12780
TGATAACATG	TCTATCAAAG	AAGCTAATGA	AACCAAAGAT	ACTAATCTTA	TTTATCAACA	12840
AATAATGTTA	ACAGGATTAA	GTGTTTCGA	ATATTTATT	AGATTAAAAG	AAACCACAGG	12900
ACACAACCC	ATAGTTATGC	ATCTGCACAT	AGAAGATGAG	TGTTGTATTA	AAGAAAGTTT	12960
TAATGATGAA	CATATTAATC	CAGAGTCTAC	ATTAGAATTA	ATTCGATATC	CTGAAAGTAA	13020
TGAATTATT	TATGATAAAG	ACCCACTCAA	AGATGTGGAC	TTATCAAAAC	TTATGGTTAT	13080
TAAAGACCAT	TCTTACACAA	TTGATATGAA	TTATTGGGAT	GATACTGACA	TCATACATGC	13140
AATTTCAATA	TGTACTGCAA	TTACAATAGC	AGATACTATG	TCACAATTAG	ATCGAGATAA	13200
TTTAAAAGAG	ATAATAGTTA	TTGCAAATGA	TGATGATATT	AATAGCTAA	TCACTGAATT	13260
TTTGACTCTT	GACATACTTG	TATTTCTCAA	GACATTTGGT	GGATTATTAG	TAAATCAATT	13320
TGCATACACT	CTTTATAGTC	AAAAAATAGA	AGGTAGGGAT	CTCATTGGG	ATTATATAAT	13380
GAGAACACTG	AGAGATACTT	CCCATTCAAT	ATTAAAAGTA	TTATCTAATG	CATTATCTCA	13440
TCCTAAAGTA	TTCAAGAGGT	TCTGGGATTG	TGGAGTTTA	AACCCTATT	ATGGTCCTAA	13500
TAUTGCTAGT	CAAGACCAGA	AAAAACTTGC	CCTATCTATA	TGTGAATATT	CACTAGATCT	13560
ATTTATGAGA	GAATGGTTGA	ATGGTGTATC	ACTTGAAATA	TACATTTGTG	ACAGCGATAT	13620
GGAAGTTGCA	AATGATAGGA	AAACAAGCCTT	TATTTCTAGA	CACCTTCAT	TTGTTGTTG	13680
TTTAGCAGAA	ATTGCATCTT	TCGGACCTAA	CCTGTTAAC	TTAACATACT	TGGAGAGACT	13740
TGATCTATTG	AAACAATATC	TTGAATTAAA	TATTAAGAA	GACCTACTC	TTAAATATGT	13800
ACAAATATCT	GGATTATTAA	TTAAATCGTT	CCCATCAACT	GTAACATACG	TAAGAAAGAC	13860
TGCAATCAA	TATCTAAGGA	TTCGCGGTAT	TAGTCCACCT	GAGGTAATTG	ATGATTGGGA	13920
TCCGGTAGAA	GATGAAAATA	TGCTGGATAA	CATTGTCAA	ACTATAAAATG	ATAACTGTAA	13980
TAAAGATAAT	AAAGGGAAATA	AAATTAACAA	TTTCTGGGGA	CTAGCACTTA	AGAACTATCA	14040
AGTCCTTAAA	ATCAGATCTA	TAACAAGTGA	TTCTGATGAT	AATGATAGAC	TAGATGCTAA	14100
TACAAGTGGT	TTGACACTTC	CTCAAGGAGG	GAATTATCTA	TCGCATCAAT	TGAGATTATT	14160
CGGAATCAAC	AGCACTAGTT	GTCTGAAAGC	TCTTGAGTTA	TCACAAATT	TAATGAAGGA	14220
AGTCAATAAA	GACAAGGACA	GGCTCTTCCT	GGGAGAAGGA	GCAGGAGCTA	TGCTAGCATG	14280
TTATGATGCC	ACATTAGGAC	CTGCAGTTAA	TTATTATAAT	TCAGGTTGA	ATATAACAGA	14340
TGTAATTGGT	CAACGAGAAT	TGAAAATATT	TCCTTCAGAG	GTATCATTAG	TAGGAAAAAA	14400
ATTAGGAAAT	GTGACACAGA	TTCTTAACAG	GGTAAAAGTA	CTGTTCAATG	GGAATCCTAA	14460
TTCAACATGG	ATAGGAAATA	TGGAATGTGA	GAGCTTAATA	TGGAGTGAAT	TAAATGATAA	14520
GTCCATTGGA	TTAGTACATT	GTGATATGGA	AGGAGCTATC	GGTAAATCAG	AAGAAACTGT	14580
TCTACATGAA	CATTATAGTG	TTATAAGAAT	TACATACTTG	ATTGGGGATG	ATGATGTTGT	14640
TTTAGTTCC	AAAATTATAC	CTACAATCAC	TCCGAATTGG	TCTAGAATAC	TTTATCTATA	14700
TAAATTATAT	TGGAAAGATG	TAAGTATAAT	ATCACTCAA	ACTTCTAATC	CTGCATCAAC	14760

AGAATTATAT CTAATTCGA AAGATGCATA TTGTACTATA ATGGAACCTA GTGAAATTGT	14820
TTTATCAAAA CTTAAAAGAT TGTCACTCTT GGAAGAAAAT AATCTATTAA AATGGATCAT	14880
TTTATCAAAG AAGAGGAATA ATGAATGGTT ACATCATGAA ATCAAAGAAG GAGAAAGAGA	14940
TTATGGAATC ATGAGACCAT ATCATATGGC ACTACAAATC TTTGGATTC AAATCAATTT	15000
AAATCATCTG CGCAAAGAAT TTTTATCAAC CCCAGATCTG ACTAATATCA ACAATATAAT	15060
CCAAAGTTT CAGCGAACAA TAAAGGATGT TTTATTTGAA TGGATTAATA TAACTCATGA	15120
TGATAAGAGA CATAAATTAG GCGGAAGATA TAACATATTC CCACTGAAAA ATAAGGGAAA	15180
GTAAAGACTG CTATCGAGAA GACTAGTATT AAGTTGGATT TCATTATCAT TATCGACTCG	15240
ATTACTTACA GGTGCGTTTC CTGATGAAAA ATTTGAACAT AGAGCACAGA CTGGATATGT	15300
ATCATTAGCT GATACTGATT TAGAACATT AAAGTTATTG TCGAAAAACA TCATTAAGAA	15360
TTACAGAGAG TGTATAGGAT CAATATCATA TTGGTTCTA ACCAAAGAAG TTAAAATACT	15420
TATGAAATTG ATCGGTGGTG CTAAATTATT AGGAATTCCC AGACAATATA AAGAACCCGA	15480
AGACCAGTTA TTAGAAAACT ACAATCAACA TGATGAATTG GATATCGATT AAAACATAAA	15540
TACAATGAAG ATATATCCTA ACCTTTATCT TTAAGCCTAG GAATAGACAA AAAGTAAGAA	15600
AAACATGTAA TATATATATA CCAACAGAG TTCTTCTCTT GTTGGTTAT AGTGAGTCGT	15660
ATTACAATC	15669

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AATACGACTC ACTATAACCA AACAAAGAGAA C

31

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCAAGTACTA TGAGATGCTT CATT

24

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CCCTATAATT TCAACATGTT GAGCCTATTT G

31

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GATTTAAATG TTGGTCGACT TAGTTGCTTC C

31

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 38 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCATAGAGAG TCCATGGAAA GCGACGCTAA AAACTATC

38

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 37 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGGTGTCGTT TCTTTGTCGA CTCATTGGCA ATTGTTG

37

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 42 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCAAAGCGTG CCCGGGCCAT GGACACTGAA TCTAACAAATG GC

42

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GAAATTCCCTT AATCGATTCT CTAGATTC

28

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCCATCAACT GTAACATACT TAAGAAAAGAC

30

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGTTAGGATA TGTCGACATT GTATTTATG

29

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 43 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGGGTTATGC TACTGCAGGC TTTTTTCTC CCTTAGCCAT CCG

43

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CTCCATTCTA GANTTATAAA AATTATAGAG TTCCC

35

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15660 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TAATACGACT CACTATAACC AAACAAGAGA AGAAACTTGT CTGGGAATAT AAATTTAACT	60
TTAAAATTAAC TTAGGATTAA AGACATTGAC TAGAAGGTCA AGAAAAGGGA ACTCTATAAT	120
TTCAAAAATG TTGAGCCTAT TTGATACATT TAATGCACGT AGGCAAGAAA ACATAACAAA	180
ATCAGCCGGT GGAGCTATCA TTCCTGGACA GAAAAATACT GTCTCTATAT TCGCCCTTGG	240
ACCGACAATA ACTGATGATA ATGAGAAAAT GACATTAGCT CTTCTATTTC TATCTCATTC	300
ACTAGATAAT GAGAAACAAAC ATGCACAAAG GGCAGGGTTC TTGGTGTCTT TATTGTCAAT	360
GGCTTATGCC AATCCAGAGC TCTACCTAAC AACAAATGGA AGTAATGCAG ATGTCAAGTA	420
TGTCATATAC ATGATTGAGA AAGATCTAAA ACGGCAAAAG TATGGAGGAT TTGTGGTTAA	480
GACGAGAGAG ATGATATATG AAAAGACAAC TGATTGGATA TTTGGAAGTG ACCTGGATTA	540
TGATCAGGAA ACTATGTTGC AGAACGGCAG GAACAATTCA ACAATTGAAG ACCTTGTCCA	600
CACATTGGG TATCCATCAT GTTTAGGAGC TCTTATAATA CAGATCTGGA TAGTTCTGGT	660
CAAAGCTATC ACTAGTATCT CAGGGTTAAG AAAAGGCTTT TTCACCCGAT TGGAAGCTTT	720
CAGACAAGAT GGAACAGTGC AGGCAGGGCT GGTATTGAGC GGTGACACAG TGGATCAGAT	780
TGGGTCAATC ATGCGGTCTC AACAGAGCTT GGTAACTCTT ATGGTTGAAA CATTAATAAC	840
AATGAATACC AGCAGAAATG ACCTCACAAAC CATAGAAAAG AATATACAAA TTGTTGGCAA	900
CTACATAAGA GATGCAGGTC TCGCTTCATT CTTCAATACA ATCAGATATG GAATTGAGAC	960
CAGAATGGCA GCTTTGACTC TATCCACTCT CAGACCAAGAT ATCAATAGAT TAAAAGCTTT	1020
GATGGAACTG TATTTATCAA AGGGACCACG CGCTCCTTTC ATCTGTATCC TCAGAGATCC	1080
TATACATGGT GAGTCGAC CAGGCAACTA TCCTGCCATA TGGAGCTATG CAATGGGGGT	1140
GGCAGTTGTA CAAAATAGAG CCATGCAACA GTATGTGACG GGAAGATCAT ATCTAGACAT	1200

TGATATGTT CAGCTAGGAC AAGCAGTAGC ACGTGATGCC	GAAGCTCAAA TGAGCTAAC	1260
ACTGGAAGAT GAACTTGGAG TGACACACGA ATCTAAAGAA	AGCTTGAAGA GACATATAAG	1320
GAACATAAAC AGTTCAGAGA CATCTTCCA CAAACCGACA	GGTGGATCAG CCATAGAGAT	1380
GGCAATAGAT GAAGAGCCAG ACAATTGCA ACATAGAGCA	GATCAAGAAC AAAATGGAGA	1440
ACCTCAATCA TCCATAATTG AATATGCCTG GGCAGAAGGA	AATAGAAGCG ATGATCAGAC	1500
TGAGCAAGCT ACAGAATCTG ACAATATCAA GACCGAACAA	CAAAACATCA GAGACAGACT	1560
AAACAAGAGA CTCAACGACA AGAAGAAACA AAGCAGTCAA	CCACCCACTA ATCCCACAAA	1620
CAGAACAAAC CAGGACGAAA TAGATGATCT GTTTAACGCA	TTTGAAGCA ACTAATCGAA	1680
TCAACATTT AATCTAAATC AATAATAAT AAGAAAAACT	TAGGATTAAA GAATCCTATC	1740
ATACCGGAAT ATAGGGTGGT AAATTAGAG TCTGCTTGAA	ACTCAATCAA TAGAGAGTTG	1800
ATGGAAAGCG ATGCTAAAAA CTATCAAATC ATGGATTCTT	GGGAAGAGGA ATCAAGAGAT	1860
AAATCAACTA ATATCTCCTC GGCCCTCAAC ATCATTGAAT	TCATACTCAG CACCGACCCC	1920
CAAGAACACT TATCGGAAAA CGACACAATC AACACAAGAA	CCCAGCAACT CAGTGCCACC	1980
ATCTGTCAAC CAGAAATCAA ACCAACAGAA ACAAGTGAGA	AAGATAGTGG ATCAACTGAC	2040
AAAAATAGAC AGTCTGGGTC ATCACACGAA TGTACAACAG	AAGCAAAAGA TAGAAATATT	2100
GATCAGGAAA CTGTACAGAG AGGACCTGGG AGAAGAAGCA	GCTCAGATAG TAGAGCTGAG	2160
ACTGTGGTCT CTGGAGGAAT CCCAGAACG ATCACAGATT	CTAAAATGG AACCCAAAAC	2220
ACGGAGGATA TTGATCTCAA TGAAATTAGA AAGATGGATA	AGGACTCTAT TGAGGGAAA	2280
ATGCGACAAT CTGCAAATGT TCCAAGCGAG ATATCAGGAA	GTGATGACAT ATTTACAACA	2340
GAACAAAGTA GAAACAGTGA TCATGGAAGA AGCCTGGAAT	CTATCAGTAC ACCTGATACA	2400
AGATCAATAA GTGTTGTTAC TGCTGCAACA CCAGATGATG	AAGAAGAAAT ACTAATGAAA	2460
AATAGTAGGA CAAAGAAAAG TTCTTCACAA CATCAAGAAG	ATGACAAAAG AATTAAAAAA	2520
GGGGGAAAAG GGAAAGACTG GTTTAAGAAA TCAAAAGATA	CCGACAACCA GATACCAACA	2580
TCAGACTACA GATCCACATC AAAAGGGCAG AAGAAAATCT	CAAAGACAAC AACCAACCAAC	2640
ACCGACACAA AGGGGCAAAC AGAAATACAG ACAGAATCAT	CAGAAACACA ATCCTCATCA	2700
TGGAATCTCA TCATCGACAA CAACACCGAC CGGAACGAAC	AGACAAGCAC AACTCCTCCA	2760
ACAACAACTT CCAGATCAAC TTATACAAAA GAATCGATCC	GAACAAACTC TGAATCCAAA	2820
CCCAAGACAC AAAAGACAAA TGAAAGGAA AGGAAGGATA	CAGAAGAGAG CAATCGATT	2880
ACAGAGAGGG CAATTACTCT ATTGCAGAAT CTTGGTGTAA	TTCAATCCAC ATCAAAACTA	2940
GATTTATATC AAGACAAACG AGTTGTATGT GTAGCAAATG	TACTAAACAA TGTAGATACT	3000
GCATCAAAGA TAGATTCCT GGCAGGATTA GTCATAGGGG	TTTCATGGGA CAACGACACA	3060
AAATTAACAC AGATACAAAA TGAAATGCTA AACCTCAAAG	CAGATCTAAA GAAAATGGAC	3120
GAATCACATA GAAGATTGAT AGAAAATCAA AGAGAACAAAC	TGTCATTGAT CACGTCACTA	3180
ATTTCAAATC TCAAAATTAT GACTGAGAGA GGAGGAAAGA	AAGACCAAAA TGAATCCAAT	3240
GAGAGAGTAT CCATGATCAA AACAAAATTG AAAGAAGAAA	AGATCAAGAA GACCAGGTTT	3300

GACCCACTTA	TGGAGGCACA	AGGCATTGAC	AAGAATATAAC	CCGATCTATA	TCGACATGCA	3360
GGAGATAACAC	TAGAGAACGA	TGTACAAGTT	AAATCAGAGA	TATTAAGTTC	ATACAATGAG	3420
TCAAATGCAA	CAAGACTAAT	ACCCAAAAAA	GTGAGCAGTA	CAATGAGATC	ACTAGTTGCA	3480
GTCATCAACA	ACAGCAATCT	CTCACAAAGC	ACAAAACAAT	CATACATAAA	CGAACTCAAA	3540
CGTTGCAAAA	ATGATGAAGA	AGTATCTGAA	TTAATGGACA	TGTTCAATGA	AGATGTCAAC	3600
AATTGCCAAT	GATCCAACAA	AGAAACGACA	CCGAACAAAC	AGACAAGAAA	CAACAGTAGA	3660
TCAAAACCTG	TCAACACACA	CAAAATCAAG	CAGAATGAAA	CAACAGATAT	CAATCAATAT	3720
ACAAATAAGA	AAAACCTAGG	ATTAAGAAT	AAATTAATCC	TTGTCAAAAA	TGAGTATAAC	3780
TAACCTTGCA	ATATACACAT	TCCCAGAATC	ATCATTCTCT	GAAAATGGTC	ATATAGAACCC	3840
ATTACCACTC	AAAGTCAATG	AACAGAGGAA	AGCAGTACCC	CACATTAGAG	TTGCCAAGAT	3900
CGGAAATCCA	CCAAAACACG	GATCCCGTA	TTTAGATGTC	TTCTTACTCG	GCTTCTTCGA	3960
GATGGAACGA	ATCAAAGACA	AATACGGGAG	TGTGAATGAT	CTCGACAGTG	ACCCGAGTTA	4020
CAAAGTTTGT	GGCTCTGGAT	CATTACCAAT	CGGATTGGCT	AAGTACACTG	GGAATGACCA	4080
GGAATTGTTA	CAAGCCGCAA	CCAAACTGGA	TATAGAAGTG	AGAAGAACAG	TCAAAGCGAA	4140
AGAGATGGTT	GTTTACACGG	TACAAAATAT	AAAACCAGAA	CTGTACCCAT	GGTCCAATAG	4200
ACTAAGAAAA	GGAATGCTGT	TCGATGCCAA	CAAAGTTGCT	CTTGCTCCTC	AATGTCTTCC	4260
ACTAGATAGG	AGCATAAAAT	TTAGAGTAAT	CTTCGTGAAT	TGTACGGCAA	TTGGATCAAT	4320
AACCTTGTTC	AAAATTCTA	AGTCAATGGC	ATCACTATCT	CTACCCAACA	CAATATCAAT	4380
CAATCTGCAG	GTACACATAA	AAACAGGGT	TCAGACTGAT	TCTAAAGGGA	TAGTTCAAAT	4440
TTTGGATGAG	AAAGGCAGAA	AATCACTGAA	TTTCATGGTC	CATCTCGGAT	TGATCAAAAG	4500
AAAAGTAGGC	AGAATGTACT	CTGTTGAATA	CTGTAAACAG	AAAATCGAGA	AAATGAGATT	4560
GATATTTCCT	TTAGGACTAG	TTGGAGGAAT	CAGTCTTCAT	GTCAATGCAA	CTGGGTCCAT	4620
ATCAAAAACA	CTAGCAAGTC	AGCTGGTATT	CAAAGAGAG	ATTGTTATC	CTTTAATGGA	4680
TCTAAATCCG	CATCTCAATC	TAGTTATCTG	GGCTTCATCA	GTAGAGATTA	CAAGAGTGGA	4740
TGCAATTTC	CAACCTTCTT	TACCTGGCGA	GTTCAGATAC	TATCCTAATA	TTATTGCAA	4800
AGGAGTTGGG	AAAATCAAAC	AATGGAACTA	GTAATCTCTA	TTTTAGTCCG	GACGTATCTA	4860
TTAAGCCGAA	GCAAATAAAG	GATAATCAA	AACTTAGGAC	AAAAGAGGTC	AATACCAACA	4920
ACTATTAGCA	GTCACACTCG	CAAGAATAAG	AGAGAAGGGA	CCAAAAAAGT	CAAATAGGAG	4980
AAATCAAAC	AAAAGGTACA	GAACACCAGA	ACAACAAAAT	CAAAACATCC	AACTCACTCA	5040
AAACAAAAAT	TCCAAAAGAG	ACCGGCAACA	CAACAAGCAC	TGAACACAAT	GCCAACTTCA	5100
ATACTGCTAA	TTATTACAAAC	CATGATCATG	GCATCTTCT	GCCAAATAGA	TATCACA	5160
CTACAGCACG	TAGGTGTATT	GGTCAACAGT	CCCAAAGGGA	TGAAGATATC	ACAAAAC	5220
GAAACAAGAT	ATCTAATT	GAGCCTCATA	CCAAAAATAG	AAGACTCTAA	CTCTTGTGGT	5280
GACCAACAGA	TCAAGCAATA	CAAGAAGTTA	TTGGATAGAC	TGATCATCCC	TTTATATGAT	5340
GGATTAAGAT	TACAGAAAGA	TGTGATAGTA	ACCAATCAAG	AATCCAATGA	AAACACTGAT	5400

CCCAGAACAA AACGATTCTT TGGAGGGTA ATTGGAACCA TTGCTCTGGG AGTAGCAACC	5460
TCAGCACAAA TTACAGCGGC AGTTGCTCTG GTTGAAGCCA AGCAGGCAAG ATCAGACATC	5520
GAAAAACTCA AAGAAGCAAT TAGGGACACA AACAAAGCAG TGCACTAGT TCAGAGCTCC	5580
ATAGGAAATT TAATAGTAGC AATTAAATCA GTCCAGGATT ATGTTAACAA AGAAATCGTG	5640
CCATCGATTG CGAGGCTAGG TTGTGAAGCA GCAGGACTTC AATTAGGAAT TGCATTAACA	5700
CAGCATTACT CAGAATTAAC AAACATATTT GGTGATAACA TAGGATCGTT ACAAGAAAAA	5760
GGAATAAAAT TACAAGGTAT AGCATCATTAC TACCGCACAA ATATCACAGA AATATTACAA	5820
ACATCAACAG TTGATAAATA TGATATCTAT GATCTGTTAT TTACAGAACATC AATAAAGGTG	5880
AGAGTTATAG ATGTTGACTT GAATGATTAC TCAATCACCC TCCAAGTCAG ACTCCCTTTA	5940
TTAACTAGGC TGCTGAACAC TCAGATCTAC AAAGTAGATT CCATATCATA TAACATCCAA	6000
AACAGAGAAT GGTATATCCC TCTTCCCAGC CATATCATGA CGAAAGGGC ATTTCTAGGT	6060
GGAGCAGACG TCAAAGAATG TATAGAAGCA TTCAGCAGCT ATATATGCC TTCTGATCCA	6120
GGATTTGTAT TAAACCATGA AATAGAGAGC TGCTTATCAG GAAACATATC CCAATGTCCA	6180
AGAACACACGG TCACATCAGA CATTGTTCCA AGATATGCAT TTGTCATGG AGGAGTGGTT	6240
GCAAACGTAA TAACAACCAC CTGTACATGC AACGGAATTG GTAATAGAAT CAATCAACCA	6300
CCTGATCAAG GAGTAAAAAT TATAACACAT AAAGAATGTA GTACAATAGG TATCAACGGA	6360
ATGCTGTTCA ATACAAATAA AGAAGGAACCTTG CATTGATTCT ATACACCAAA TGATATAACA	6420
CTAAACAATT CTGTTGCACT TGATCCAATT GACATATCAA TCGAGCTCAA CAAGGCCAAA	6480
TCAGATCTAG AAGAATCAA AGAATGGATA AGAAGGTCAA ATCAAAACT AGATTCTATT	6540
GGAAATTGGC ATCAATCTAG CACTACAATC ATAATTATTT TGATAATGAT CATTATATTG	6600
TTTATAATTA ATATAACGAT AATTACAATT GCAATTAAGT ATTACAGAAT TCAAAAGAGA	6660
AATCGAGTGG ATCAAAATGA CAAGCCATAT GTACTAACAA ACAAAATAACA TATCTACAGA	6720
TCATTAGATA TTAAAATTAT AAAAAGCTTA GGAGTAAAGT TACGCAATCC AACTCTACTC	6780
ATATAATTGA GGAAGGACCC AATAGACAAA TCCAAATTGAG AGATGGAATA CTGGAAGCAT	6840
ACCAATCACG GAAAGGATGC TGGTAATGAG CTGGAGACGT CTATGGCTAC TCATGGCAAC	6900
AAGCTCACTA ATAAGATAAT ATACATATTA TGGACAATAA TCCTGGTGT ATTATCAATA	6960
GTCTTCATCA TAGTGCTAAT TAATTCCATC AAAAGTGAAA AGGCCACGA ATCATTGCTG	7020
CAAGACATAA ATAATGAGTT TATGGAAATT ACAGAAAAGA TCCAAATGGC ATCGGATAAT	7080
ACCAATGATC TAATACAGTC AGGAGTGAAT ACAAGGCTTC TTACAATTCA GAGTCATGTC	7140
CAGAATTACA TACCAATATC ATTGACACAA CAGATGTCAG ATCTTAGGAA ATTCTTAAAGT	7200
GAAATTACAA TTAGAAATGA TAATCAAGAA GTGCTGCCAC AAAGAATAAC ACATGATGTA	7260
GGTATAAAAC CTTTAAATCC AGATGATTT TGGAGATGCA CGTCTGGTCT TCCATCTTTA	7320
ATGAAAACCTC CAAAATAAG GTTAATGCCA GGGCCGGGAT TATTAGCTAT GCCAACGACT	7380
GTTGATGGCT GTGTTAGAAC TCCGTCTTTA GTTATAAAATG ATCTGATTAA TGCTTATACC	7440
TCAAATCTAA TTACTCGAGG TTGTCAGGAT ATAGGAAAAT CATATCAAGT CTTACAGATA	7500

GGGATAATAA CTGTAAACTC AGACTTGGTA CCTGACTTAA ATCCTAGGAT CTCTCATACC	7560
TTTAACATAA ATGACAATAG GAAGTCATGT TCTCTAGCAC TCCTAAATAT AGATGTATAT	7620
CAACTGTGTT CAACTCCCAA AGTTGATGAA AGATCAGATT ATGCATCATC AGGCATAGAA	7680
GATATTGTAC TTGATATTGT CAATTATGAT GGTTCAATCT CAACAACAAG ATTTAAGAAT	7740
AATAACATAA GCTTTGATCA ACCATATGCT GCACTATACC CATCTGTTGG ACCAGGGATA	7800
TACTACAAAG GCAAAATAAT ATTTCTCGGG TATGGAGGTC TTGAACATCC AATAAATGAG	7860
AATGTAATCT GCAACACAAAC TGGGTGCCCG GGGAAAACAC AGAGAGACTG TAATCAAGCA	7920
TCTCATAGTA CTTGGTTTC AGATAGGAGG ATGGTCAACT CCATCATTGT TGTTGACAAA	7980
GGCTTAAACT CAATTCCAAA ATTGAAAGTA TGGACGATAT CTATGCGACA AAATTACTGG	8040
GGGTCAGAAG GAAGGTTACT TCTACTAGGT AACAAAGATCT ATATATATAC AAGATCTACA	8100
AGTTGGCATA GCAAGTTACA ATTAGGAATA ATTGATATTA CTGATTACAG TGATATAAGG	8160
ATAAAATGGA CATGGCATAA TGTGCTATCA AGACCAGGAA ACAATGAATG TCCATGGGGA	8220
CATTCATGTC CAGATGGATG TATAACAGGA GTATATACTG ATGCATATCC ACTCAATCCC	8280
ACAGGGAGCA TTGTGTCATC TGTATATTAA GACTCACAAA AATCGAGAGT GAACCCAGTC	8340
ATAACTTACT CAACAGCAAC CGAAAGAGTA AACGAGCTGG CCATCCTAAA CAGAACACTC	8400
TCAGCTGGAT ATACAACAAAC AAGCTGCATT ACACACTATA ACAAAAGGATA TTGTTTCAT	8460
ATAGTAGAAA TAAATCATAA AAGCTTAAAC ACATTTCAAC CCATGTTGTT CAAAACAGAG	8520
ATTCCAAAAA GCTGCAGTTA ATCATAATTAA ACCATAATAT GCATCAATCT ATCTATAATA	8580
CAAGTATATG ATAAGTAATC AGCAATCAGA CAATAGACAA AAGGGAAATA TAAAAAAACTT	8640
AGGAGCAAAG CGTGCTCGGG AAATGGACAC TGAATCTAAC AATGGCACTG TATCTGACAT	8700
ACTCTATCCT GAGTGTCAACC TTAACTCTCC TATCGTTAAA GGTTAAATAG CACAATTACA	8760
CACTATTATG AGTCTACCTC AGCCTTATGA TATGGATGAC GACTCAATAC TAGTTATCAC	8820
TAGACAGAAA ATAAAACCTTA ATAAATTGGA TAAAAGACAA CGATCTATTAA GAAGATTAAA	8880
ATTAATATTAA ACTGAAAAAG TGAATGACTT AGGAAAATAC ACATTTATCA GATATCCAGA	8940
AATGTCAAAA GAAATGTTCA AATTATATAT ACCTGGTATT AACAGTAAAG TGACTGAATT	9000
ATTACTTAAA GCAGATAGAA CATATAGTC AATGACTGAT GGATTAAGAG ATCTATGGAT	9060
TAATGTGCTA TCAAAATTAG CCTCAAAAAA TGATGGAAGC AATTATGATC TTAATGAAGA	9120
AATTAATAAT ATATCGAAAG TTCACACAAAC CTATAATCA GATAATGGT ATAATCCATT	9180
CAAAACATGG TTTACTATCA AGTATGATAT GAGAAGATTA CAAAAAGCTC GAAATGAGAT	9240
CACTTTAAT GTTGGGAAGG ATTATAACTT GTTAGAAGAC CAGAAGAATT TCTTATTGAT	9300
ACATCCAGAA TTGGTTTGA TATTAGATAA ACAAAACATAT AATGGTTATC TAATTACTCC	9360
TGAATTAGTA TTGATGTATT GTGACGTAGT CGAAGGCCGA TGGAATATAA GTGCATGTGC	9420
TAAGTTAGAT CCAAAATTAC AATCTATGTA TCAGAAAGGT AATAACCTGT GGGAAAGTGAT	9480
AGATAAAATTG TTTCCAATTAA TGGGAGAAAA GACATTTGAT GTGATATCGT TATTAGAAC	9540
ACTTGCATTA TCCTTAATTCAAACTCATGA TCCTGTTAAA CAACTAAGAG GAGCTTTTT	9600

AAATCATGTG TTATCCGAGA TGGAATTAAT ATTTGAATCT AGAGAACATCA TTAAGGAATT	9660
TCTGAGTGTA GATTACATTG ATAAAATTTT AGATATATTT AATAAGTCTA CAATAGATGA	9720
AATAGCAGAG ATTTCTCTT TTTTAGAAC ATTTGGGCAT CCTCCATTAG AAGCTAGTAT	9780
TGCAGCAGAA AAGGTTAGAA AATATATGTA TATTGGAAAA CAATTAAAT TTGACACTAT	9840
TAATAAATGT CATGCTATCT TCTGTACAAT AATAATTAAC GGATATAGAG AGAGGCATGG	9900
TGGACAGTGG CCTCCTGTGA CATTACCTGA TCATGCACAC GAATTCATCA TAAATGCTTA	9960
CGGTTCAAAC TCTGCGATAT CATATGAAAA TGCTGTTGAT TATTACCAGA GCTTTATAGG	10020
AATAAAATTC AATAAATTCA TAGAGCCTCA GTTAGATGAG GATTTGACAA TTTATATGAA	10080
AGATAAAAGCA TTATCTCCAA AAAAATCAA TTGGGACACA GTTTATCCTG CATCTAATTT	10140
ACTGTACCGT ACTAACGCAT CCAACGAATC ACGAAGATTA GTTGAAGTAT TTATAGCAGA	10200
TAGTAAATTT GATCCTCATC AGATATTGGA TTATGTAGAA TCTGGGACT GGTTAGATGA	10260
TCCAGAATTT AATATTTCTT ATAGTCTTAA AGAAAAGAG ATCAAACAGG AAGGTAGACT	10320
CTTGCAAAA ATGACATACA AAATGAGAGC TACACAAGTT TTATCAGAGA CCCTACTTGC	10380
AAATAACATA GGAAAATTCT TTCAAGAAAA TGGGATGGTG AAGGGAGAGA TTGAATTACT	10440
TAAGAGATTA ACAACCATA CAATATCAGG AGTTCCACGG TATAATGAAG TGTACAATAA	10500
TTCTAAAAGC CATACTAGATG ACCTTAAAC CTACAATAAA ATAAGTAATC TTAATTGTC	10560
TTCTAAATCAG AAATCAAAGA AATTGAAATT CAAGTCACG GATATCTACA ATGATGGATA	10620
CGAGACTGTG AGCTGTTCC TAACAACAGA TCTCAAAAAA TACTGTCTTA ATTGGAGATA	10680
TGAATCAACA GCTCTATTG GAGAAACTTG CAACCAAATA TTTGGATTAA ATAAATTGTT	10740
TAATTGGTTA CACCCTCGTC TTGAAGGAAG TACAATCTAT GTAGGTGATC CTTACTGTCC	10800
TCCATCAGAT AAAGAACATA TATCATTAGA GGATCACCT GATTCTGGTT TTTACGTTCA	10860
TAACCCAAGA GGGGGTATAG AAGGATTTG TCAAAAATTA TGGACACTCA TATCTATAAG	10920
TGCAATACAT CTAGCAGCTG TTAGAATAGG CGTGAGGGTG ACTGCAATGG TTCAAGGAGA	10980
CAATCAAGCT ATAGCTGTAA CCACAAGAGT ACCCAACAAT TATGACTACA GAGTTAAGAA	11040
GGAGATAGTT TATAAAGATG TAGTGAGATT TTTTGATTCA TTAAGAGAAG TGATGGATGA	11100
TCTAGGTCAAT GAACTTAAAT TAAATGAAAC GATTATAAGT AGCAAGATGT TCATATATAG	11160
CAAAAGAAC TATTATGATG GGAGAATTCT TCCTCAAGCT CTAAAAGCAT TATCTAGATG	11220
TGTCTTCTGG TCAGAGACAG TAATAGACGA AACAAAGATCA GCATCTCAA ATTTGGCAAC	11280
ATCATTGCA AAAGCAATTG AGAATGGTTA TTCACCTGTT CTAGGATATG CATGCTCAAT	11340
TTTTAAGAAT ATTCAACAAAC TATATATTGC CCTTGGGATG AATATCAATC CAACTATAAC	11400
ACAGAACATC AGAGATCAGT ATTTAGGAA TCCAAATTGG ATGCAATATG CCTCTTTAAT	11460
ACCTGCTAGT GTTGGGGGAT TCAATTACAT GGCCATGTCA AGATGTTTG TAAGGAATAT	11520
TGGTGATCCA TCAGTTGCCG CATTGGCTGA TATTAAAAGA TTTATTAAGG CGAATCTATT	11580
AGACCGAAGT GTTCTTTATA GGATTATGAA TCAAGAACCA GGTGAGTCAT CTTTTTGGA	11640
CTGGGCTTCA GATCCATATT CATGCAATTT ACCACAATCT CAAAATATAA CCACCATGAT	11700

AAAAAAATATA ACAGCAAGGA ATGTATTACA AGATTCACCA AATCCATTAT TATCTGGATT	11760
ATTCACAAAT ACAATGATAG AAGAAGATGA AGAATTAGCT GAGTTCCTGA TGGACAGGAA	11820
GGTAATTCTC CCTAGAGTTG CACATGATAT TCTAGATAAT TCTCTCACAG GAATTAGAAA	11880
TGCCATAGCT GGAATGTTAG ATACGACAAA ATCACTAATT CGGGTTGGCA TAAATAGAGG	11940
AGGACTGACA TATAGTTGT TGAGGAAAAT CAGTAATTAC GATCTAGTAC AATATGAAAC	12000
ACTAAGTAGG ACTTTGCGAC TAATTGTAAG TGATAAAATC AAGTATGAAG ATATGTGTTC	12060
GGTAGACCTT GCCATAGCAT TGCGACAAAA GATGTGGATT CATTATCAG GAGGAAGGAT	12120
GATAAGTGGA CTTGAAACGC CTGACCCATT AGAATTACTA TCTGGGGTAG TAATAACAGG	12180
ATCAGAACAT TGTAAAATAT GTTATTCTTC AGATGGCACA AACCCATATA CTTGGATGTA	12240
TTTACCCGGT AATATCAAAA TAGGATCAGC AGAAACAGGT ATATCGTCAT TAAGAGTTCC	12300
TTATTTGGA TCAGTCACTG ATGAAAGATC TGAAGCACAA TTAGGATATA TCAAGAATCT	12360
TAGTAAACCT GCAAAAGCCG CAATAAGAAT AGCAATGATA TATACATGGG CATTGGTAA	12420
TGATGAGATA TCTGGATGG AAGCCTCACA GATAGCACAA ACACGTGCAA ATTTTACACT	12480
AGATAGTCTC AAAATTTAA CACCGGTAGC TACATCAACA AATTTATCAC ACAGATTAAA	12540
GGATACTGCA ACTCAGATGA AATTCTCCAG TACATCATTG ATCAGAGTCA GCAGATTCAT	12600
AACAATGTCC AATGATAACA TGTCTATCAA AGAAGCTAAT GAAACCAAAG ATACTAATCT	12660
TATTTATCAA CAAATAATGT TAACAGGATT AAGTGTTC GAATATTTAT TTAGATTAAA	12720
AGAAAACCACA GGACACAACC CTATAGTTAT GCATCTGCAC ATAGAAGATG AGTGGTGTAT	12780
TAAAGAAAGT TTAATGATG AACATATTAA TCCAGAGTCT ACATTAGAAT TAATTGATA	12840
TCCTGAAAGT AATGAATTAA TTTATGATAA AGACCCACTC AAAGATGTGG ACTTATCAA	12900
ACTTATGGTT ATTAAAGACC ATTCTTACAC AATTGATATG AATTATTGGG ATGATACTGA	12960
CATCATAACAT GCAATTCAA TATGACTGC AATTACAATA GCAGATACTA TGTCACAATT	13020
AGATCGAGAT AATTTAAAG AGATAATAGT TATTGCAAAT GATGATGATA TTAATAGCTT	13080
AATCACTGAA TTTTGACTC TTGACATACT TGTATTTCTC AAGACATTG GTGGATTATT	13140
AGTAAATCAA TTTGCATACA CTCTTATAG TCTAAAATA GAAGGTAGGG ATCTCATTG	13200
GGATTATATA ATGAGAACAC TGAGAGATAC TTCCCATCA ATATTAAAG TATTATCTAA	13260
TGCATTATCT CATCCTAAAG TATTCAAGAG GTTCTGGAT TGTGGAGTTT TAAACCTTAT	13320
TTATGGCCT AATACTGCTA GTCAAGACCA GATAAAACTT GCCCTATCTA TATGTGAATA	13380
TTCACTAGAT CTATTTATGA GAGAATGGTT GAATGGTGT A CACTTGAAA TATACATTG	13440
TGACAGCGAT ATGGAAGTTG CAAATGATAG GAAACAAAGCC TTTATTTCTA GACACCTTTC	13500
ATTTGTTGT TGTGAGAGA AAATTGCATC TTTCGGACCT AACCTGTTAA ACTTAAACATA	13560
CTTGGAGAGA CTTGATCTAT TGAAACAATA TCTTGAATTA AATATTAAAG AAGACCCTAC	13620
TCTTAAATAT GTACAAATAT CTGGATTATT AATTAAATCG TTCCCATCAA CTGTAACATA	13680
CGTAAGAAAG ACTGCAATCA AATATCTAAG GATTGCGGT ATTGTCACAC CTGAGGTAAT	13740
TGATGATTGG GATCCGGTAG AAGATGAAAA TATGCTGGAT AACATTGTCA AACTATAAA	13800

TGATAACTGT AATAAAGATA ATAAAGGGAA TAAAATTAAC AATTCTGGG GACTAGCACT	13860
TAAGAACTAT CAAGTCCTTA AAATCAGATC TATAACAAGT GATTCTGATG ATAATGATAG	13920
ACTAGATGCT AATACAAGTG GTTGACACT TCCTCAAGGA GGGATTATC TATCGCATCA	13980
ATTGAGATTA TTCGGAATCA ACAGCACTAG TTGTCTGAAA GCTCTTGAGT TATCACAAAT	14040
TTTAATGAAG GAAGTCAATA AAGACAAGGA CAGGCTCTTC CTGGGAGAAG GAGCAGGAGC	14100
TATGCTAGCA TGTTATGATG CCACATTAGG ACCTGCAGTT AATTATTATA ATTCAAGTTT	14160
GAATATAACA GATGTAATTG GTCAACGAGA ATTGAAAATA TTTCCCTTCAG AGGTATCATT	14220
AGTAGGTAAA AAATTAGGAA ATGTGACACA GATTCTAAC AGGGTAAAG TACTGTTCAA	14280
TGGGAATCCT AATTCAACAT GGATAGGAAA TATGGAATGT GAGAGCTTAA TATGGAGTGA	14340
ATTAAATGAT AAGTCCATTG GATTAGTACA TTGTGATATG GAAGGAGCTA TCGGTAAATC	14400
AGAAGAAACT GTTCTACATG AACATTATAG TGTTATAAGA ATTACATACT TGATTGGGA	14460
TGATGATGTT GTTTTAGTTT CAAAAATTAT ACCTACAATC ACTCCGAATT GGTCTAGAAT	14520
ACTTTATCTA TATAAATTAT ATTGAAAGA TGTAAGTATA ATATCACTCA AAACTTCTAA	14580
TCCTGCATCA ACAGAATTAT ATCTAATTTC GAAAGATGCA TATTGTACTA TAATGGAACC	14640
TAGTGAAATT GTTTTATCAA AACTAAAAG ATTGTCACTC TTGGAAGAAA ATAATCTATT	14700
AAAATGGATC ATTTTATCAA AGAAGAGGAA TAATGAATGG TTACATCATG AAATCAAAGA	14760
AGGAGAAAGA GATTATGGAA TCATGAGACC ATATCATATG GCACTACAAA TCTTTGGATT	14820
TCAAATCAAT TAAATCATC TGGCGAAAGA ATTTTATCA ACCCCAGATC TGACTAATAT	14880
CAACAATATA ATCCAAAGTT TTCAGCGAAC AATAAAGGAT GTTTTATTTG AATGGATTAA	14940
TATAACTCAT GATGATAAGA GACATAAATT AGGCGGAAGA TATAACATAT TCCCAC TGAA	15000
AAATAAGGGA AAGTTAACAGC TGCTATCGAG AAGACTAGTA TTAAGTTGGA TTTCATTATC	15060
ATTATCGACT CGATTACTTA CAGGTCGCTT TCCTGATGAA AAATTTGAAC ATAGAGCACA	15120
GACTGGATAT GTATCATTAG CTGATACTGA TTTAGAATCA TTAAAGTTAT TGTCGAAAAAA	15180
CATCATTAAG AATTACAGAG AGTGTATAGG ATCAATATCA TATTGGTTTC TAACCAAAGA	15240
AGTTAAAATA CTTATGAAAT TGATCGGTGG TGCTAAATTA TTAGGAATTC CCAGACAATA	15300
TAAAGAACCC GAAGACCAGT TATTAGAAAA CTACAATCAA CATGATGAAT TTGATATCGA	15360
TTAAAACATA AATACAATGA AGATATATCC TAACCTTAT CTTAAGCCT AGGAATAGAC	15420
AAAAAGTAAG AAAAACATGT AATATATATA TACCAAACAG AGTCTTCTC TTGTTGGTG	15480
GGTCGGCATG GCATCTCCAC CTCCTCGCGG TCCGGACCTG GGCATCCGAA GGAGGACGCA	15540
CGTCCACTCG GATGGCTAAG GGAGAGCCTG CAGTAGCATA ACCCCTTGGG GCCTCTAAC	15600
GGGTCTTGAG GGGTTTTTG CTGAAAGGAG GAACTATATA CGCGTCGACG GGCCCCGCGC	15660

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15666 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TAATACGACT CACTATAGGA CCAAACAAGA GAAGAAACTT GTCTGGGAAT ATAAATTAA	60
CTTTAAATTA ACTTAGGATT AAAGACATTG ACTAGAAGGT CAAGAAAAGG GAACTCTATA	120
ATTTCAAAAA TGTTGAGCCT ATTTGATACA TTTAATGCAC GTAGGCAAGA AACATATAACA	180
AAATCAGCCG GTGGAGCTAT CATTCCCTGGA CAGAAAAATA CTGTCTCTAT ATTGCCCTT	240
GGACCGACAA TAACTGATGA TAATGAGAAA ATGACATTAG CTCTTCTATT TCTATCTCAT	300
TCACTAGATA ATGAGAAACA ACATGCACAA AGGGCAGGGT TCTTGGTGTCA TTTATTGTCA	360
ATGGCTTATG CCAATCCAGA GCTCTACCTA ACAACAAATG GAAGTAATGC AGATGTCAAG	420
TATGTCATAT ACATGATTGA GAAAGATCTA AAACGGCAAA AGTATGGAGG ATTTGTGGTT	480
AAGACGAGAG AGATGATATA TGAAAAGACA ACTGATTGGA TATTGGAAG TGACCTGGAT	540
TATGATCAGG AAACTATGTT GCAGAACGGC AGGAACAATT CAACAATTGA AGACCTTGTC	600
CACACATTG GGTATCCATC ATGTTAGGA GCTCTTATAA TACAGATCTG GATAGTTCTG	660
GTCAAAGCTA TCACTAGTAT CTCAGGGTTA AGAAAAGGCT TTTCACCCG ATTGGAAGCT	720
TTCAGACAAG ATGGAACAGT GCAGGCAGGG CTGGTATTGA GCGGTGACAC AGTGGATCAG	780
ATTGGGTCAA TCATGCGGTC TCAACAGAGC TTGGTAACCTC TTATGGTTGA AACATTAATA	840
ACAATGAATA CCAGCAGAAA TGACCTCACA ACCATAGAAA AGAATATACA AATTGTTGGC	900
AACTACATAA GAGATGCAGG TCTCGCTTCA TTCTTCAATA CAATCAGATA TGGAATTGAG	960
ACCAGAATGG CAGCTTTGAC TCTATCCACT CTCAGACCAAG ATATCAATAG ATTAAAAGCT	1020
TTGATGGAAC TGTATTTATC AAAGGGACCA CGCGCTCCTT TCATCTGTAT CCTCAGAGAT	1080
CCTATACATG GTGAGTTCGC ACCAGGCAAC TATCCTGCCA TATGGAGCTA TGCAATGGGG	1140
GTGGCAGTTG TACAAAATAG AGCCATGCAA CAGTATGTGA CGGGAAGATC ATATCTAGAC	1200
ATTGATATGT TCCAGCTAGG ACAAGCAGTA GCACGTGATG CCGAAGCTCA AATGAGCTCA	1260
ACACTGGAAG ATGAACTTGG AGTGACACAC GAATCTAAAG AAAGCTTGAA GAGACATATA	1320
AGGAACATAA ACAGTTCAGA GACATCTTC CACAAACCGA CAGGTGGATC AGCCATAGAG	1380
ATGGCAATAG ATGAAGAGCC AGAACAAATTC GAACATAGAG CAGATCAAGA ACAAAATGGA	1440
GAACCTCAAT CATCCATAAT TCAATATGCC TGGGCAGAAG GAAATAGAAG CGATGATCAG	1500
ACTGAGCAAG CTACAGAACATC TGACAATATC AAGACCGAAC AACAAACAT CAGAGACAGA	1560
CTAAACAAGA GACTAACGCA CAAGAAGAAA CAAAGCAGTC AACCAACCCAC TAATCCCACA	1620
AACAGAACAA ACCAGGACGA AATAGATGAT CTGTTAACG CATTGGAAG CAACTAATCG	1680
AATCAACATT TTAATCTAAA TCAATAATAA ATAAGAAAAA CTTAGGATTA AAGAACCTA	1740
TCATACCGGA ATATAGGGTG GTAAATTTAG AGTCTGCTTG AAACTCAATC AATAGAGAGT	1800
TGATGGAAG CGATGCTAAA AACTATCAA TCATGGATTC TTGGGAAGAG GAATCAAGAG	1860

ATAAATCAAC TAATATCTCC TCGGCCCTCA ACATCATTGA ATTCTACTC AGCACCGACC	1920
CCCAAGAAGA CTTATCGGAA AACGACACAA TCAACACAAG AACCCAGCAA CTCAGTGCCTA	1980
CCATCTGTCA ACCAGAAATC AAACCAACAG AAACAAGTGA GAAAGATAGT GGATCAACTG	2040
ACAAAAATAG ACAGTCTGGG TCATCACACG AATGTACAAC AGAAGCAAAA GATAGAAATA	2100
TTGATCAGGA AACTGTACAG AGAGGACCTG GGAGAAGAAG CAGCTCAGAT AGTAGAGCTG	2160
AGACTGTGGT CTCTGGAGGA ATCCCCAGAA GCATCACAGA TTCTAAAAAT GGAACCCAAA	2220
ACACGGAGGA TATTGATCTC AATGAAATTAA GAAAGATGGA TAAGGACTCT ATTGAGGGGA	2280
AAATGCGACA ATCTGCAAAT GTTCCAAGCG AGATATCAGG AAGTGATGAC ATATTTACAA	2340
CAGAACAAAG TAGAAACAGT GATCATGGAA GAAGCCTGGA ATCTATCAGT ACACCTGATA	2400
CAAGATCAAT AAGTGTGTT ACTGCTGCAA CACCAGATGA TGAAGAAGAA ATACTAATGA	2460
AAAATAGTAG GACAAAGAAA AGTTCTCAA CACATCAAGA AGATGACAAA AGAATTAAAA	2520
AAGGGGGAAA AGGGAAAGAC TGGTTTAAGA AATCAAAAGA TACCGACAAC CAGATACCAA	2580
CATCAGACTA CAGATCCACA TCAAAAGGGC AGAAGAAAAT CTCAAAGACA ACAACCACCA	2640
ACACCGACAC AAAGGGGCAA ACAGAAATAC AGACAGAATC ATCAGAAACA CAATCCTCAT	2700
CATGGAATCT CATCATCGAC AACAAACACCG ACCGGAACGA ACAGACAAGC ACAACTCCTC	2760
CAACAACAAAC TTCCAGATCA ACTTATACAA AAGAATCGAT CCGAACAAAC TCTGAATCCA	2820
AACCCAAGAC ACAAAAGACA AATGGAAAGG AAAGGAAGGA TACAGAAGAG AGCAATCGAT	2880
TTACAGAGAG GGCAATTACT CTATTGCAGA ATCTTGGTGT AATTCAATCC ACATCAAAAC	2940
TAGATTATA TCAAGACAAA CGAGTTGTAT GTGTAGCAAA TGTACTAAC AATGTAGATA	3000
CTGCATCAA GATAGATTTC CTGGCAGGAT TAGTCATAGG GGTTCAATG GACAACGACA	3060
CAAAATTAAAC ACAGATACAA AATGAAATGC TAAACCTCAA AGCAGATCTA AAGAAAATGG	3120
ACGAATCACA TAGAAGATTG ATAGAAAATC AAAGAGAACAA ACTGTCATTG ATCACGTCAC	3180
TAATTTCAAAT TCTCAAAATT ATGACTGAGA GAGGAGGAAA GAAAGACCAA AATGAATCCA	3240
ATGAGAGAGT ATCCATGATC AAAACAAAAT TGAAAGAAGA AAAGATCAAG AAGACCAGGT	3300
TTGACCCACT TATGGAGGCA CAAGGCATTG ACAAGAATAT ACCCGATCTA TATCGACATG	3360
CAGGAGATAC ACTAGAGAAC GATGTACAAG TTAAATCAGA GATATTAAGT TCATACAAATG	3420
AGTCAAATGC ACAAAAGACTA ATACCCAAAA AAGTGAGCAG TACAATGAGA TCACACTTG	3480
CAGTCATCAA CAACAGCAAT CTCTCACAAA GCACAAAACA ATCATACATA AACGAACCTCA	3540
AACGTTGCAA AAATGATGAA GAAGTATCTG AATTAATGGA CATGTTCAAT GAAGATGTCA	3600
ACAATTGCCA ATGATCCAAC AAAGAAACGA CACCGAACAA ACAGACAAGA AACAAACAGTA	3660
GATCAAAACC TGTCAACACA CACAAAATCA AGCAGAATGA AACAAACAGAT ATCAATCAAT	3720
ATACAAATAA GAAAAACTTA GGATTAAGA ATAAATTAAT CCTTGTCCAA AATGAGTATA	3780
ACTAACTCTG CAATATACAC ATTCCCAGAA TCATCATTCT CTGAAAATGG TCATATAGAA	3840
CCATTACAC TCAAAGTCAGA TGAACAGAGG AAAGCAGTAC CCCACATTAG AGTTGCCAAG	3900
ATCGGAAATC CACCAAAACA CGGATCCCAGG TATTTAGATG TCTTCTTACT CGGCTTCTTC	3960

GAGATGGAAC GAATCAAAGA CAAATACGGG AGTGTGAATG ATCTCGACAG TGACCCGAGT	4020
TACAAAGTTT GTGGCTCTGG ATCATTACCA ATCGGATTGG CTAAGTACAC TGGGAATGAC	4080
CAGGAATTGT TACAAGCCGC AACCAAACGT GATATAGAAC TGAGAAGAAC AGTCAAAGCG	4140
AAAGAGATGG TTGTTTACAC GGTACAAAAT ATAAAACCAG AACTGTACCC ATGGTCCAAT	4200
AGACTAAGAA AAGGAATGCT GTTCGATGCC AACAAAGTTG CTCTTGCTCC TCAATGTCTT	4260
CCACTAGATA GGAGCATAAA ATTTAGAGTA ATCTTCGTGA ATTGTACGGC AATTGGATCA	4320
ATAAACCTTGT TCAAAATTCC TAAGTCAATG GCATCACTAT CTCTACCCAA CACAATATCA	4380
ATCAATCTGC AGGTACACAT AAAAACAGGG GTTCAGACTG ATTCTAAAGG GATAGTTCAA	4440
ATTTTGGATG AGAAAGGCGA AAAATCACTG AATTCATGG TCCATCTCGG ATTGATCAA	4500
AGAAAAGTAG GCAGAATGTA CTCTGTTGAA TACTGTAAAC AGAAAATCGA GAAAATGAGA	4560
TTGATATTTT CTTTAGGACT AGTTGGAGGA ATCAGTCTTC ATGTCAATGC AACTGGTCC	4620
ATATCAAAAA CACTAGCAAG TCAGCTGGTA TTCAAAAGAG AGATTTGTTA TCCTTTAATG	4680
GATCTAAATC CGCATCTCAA TCTAGTTATC TGGGCTTCAT CAGTAGAGAT TACAAGAGTG	4740
GATGCAATTT TCCAACCTTC TTTACCTGGC GAGTTCAGAT ACTATCCTAA TATTATTGCA	4800
AAAGGAGTTG GGAAAATCAA ACAATGGAAC TAGTAATCTC TATTTTAGTC CGGACGTATC	4860
TATTAAGCCG AAGCAAATAA AGGATAATCA AAAACTTAGG ACAAAAGAGG TCAATACCAA	4920
CAACTATTAG CAGTCACACT CGCAAGAATA AGAGAGAAGG GACCAAAAAA GTCAAATAGG	4980
AGAAAATCAA ACAAAAGGTA CAGAACACCA GAACAACAAA ATCAAAACAT CCAACTCACT	5040
CAAAACAAAA ATTCCAAAAG AGACCGGCAA CACAACAAGC ACTGAACACCA ATGCCAACTT	5100
CAAAACTGCT AATTATTACA ACCATGATCA TGGCATCTT CTGCCAAATA GATATCACAA	5160
AACTACAGCA CGTAGGTGTA TTGGTCAACA GTCCCAAAGG GATGAAGATA TCACAAAATC	5220
TTGAAACAAG ATATCTAATT TTGAGCCTCA TACAAAAAT AGAAGACTCT AACTCTGTG	5280
GTGACCAACA GATCAAGCAA TACAAGAAGT TATTGGATAG ACTGATCATC CCTTTATATG	5340
ATGGATTAAG ATTACAGAAA GATGTGATAG TAACCAATCA AGAATCCAAT GAAAACACTG	5400
ATCCCAGAAC AAAACGATTG TTTGGAGGGG TAATTGGAAC CATTGCTCTG GGAGTAGCAA	5460
CCTCAGCAC AATTACAGCG GCAGTTGCTC TGGTTGAAGC CAAGCAGGCA AGATCAGACA	5520
TCGAAAAACT CAAAGAAGCA ATTAGGGACA CAAACAAAGC AGTGCAGTCA GTTCAGAGCT	5580
CCATAGGAAA TTTAATAGTA GCAATTAAAT CAGTCCAGGA TTATGTTAAC AAAGAAATCG	5640
TGCCATCGAT TCGAGGGCTA GGTTGTGAAG CAGCAGGACT TCAATTAGGA ATTGCATTAA	5700
CACAGCATTAA CTCAGAAATTA ACAAAACATAT TTGGTGATAA CATAGGATCG TTACAAGAAA	5760
AAGGAATAAA ATTACAAGGT ATAGCATCAT TATACCGCAC AAATATCACA GAAATATTCA	5820
CAACATCAAC AGTTGATAAA TATGATATCT ATGATCTGTT ATTTACAGAA TCAATAAAGG	5880
TGAGAGTTAT AGATGTTGAC TTGAATGATT ACTCAATCAC CCTCCAAGTC AGACTCCCTT	5940
TATTAACTAG GCTGCTGAAC ACTCAGATCT ACAAAAGTAGA TTCCATATCA TATAACATCC	6000
AAAACAGAGA ATGGTATATC CCTCTTCCCA GCCATATCAT GACGAAAGGG GCATTCTAG	6060

GTGGAGCAGA CGTCAAAGAA TGTATAGAAG CATTCAAGCAG CTATATATGC CCTTCTGATC	6120
CAGGATTGT ATTAAACCAT GAAATAGAGA GCTGCTTATC AGGAAACATA TCCCAATGTC	6180
CAAGAACAAAC GGTACACATCA GACATTGTTCA CAAGATATGC ATTTGTCAAT GGAGGAGTGG	6240
TTGCAAACGT TATAACAACC ACCTGTACAT GCAACGGAAT TGGTAATAGA ATCAATCAAC	6300
CACCTGATCA AGGAGTAAAA ATTATAACAC ATAAAGAATG TAGTACAATA GGTATCAACG	6360
GAATGCTGTT CAATACAAAT AAAGAAGGAA CTCTTGCTT CTATACACCA AATGATATAA	6420
CACTAAACAA TTCTGTTGCA CTTGATCCAA TTGACATATC AATCGAGCTC AACAAAGGCCA	6480
AATCAGATCT AGAAGAATCA AAAGAATGGA TAAGAAGGTC AAATCAAAA CTAGATTCTA	6540
TTGGAAATTG GCATCAATCT AGCACTACAA TCATAATTAT TTTGATAATG ATCATTATAT	6600
TGTTTATAAT TAATATAACG ATAATTACAA TTGCAATTAA GTATTACAGA ATTCAAAAGA	6660
GAAATCGAGT GGATCAAAAT GACAAGCCAT ATGTACTAAC AAACAAATAA CATATCTACA	6720
GATCATTAGA TATTAAAATT ATAAAAAACT TAGGAGTAAA GTTACGCAAT CCAACTCTAC	6780
TCATATAATT GAGGAAGGAC CCAATAGACA AATCCAAATT CGAGATGGAA TACTGGAAGC	6840
ATACCAATCA CGGAAAGGAT GCTGGTAATG AGCTGGAGAC GTCTATGGCT ACTCATGGCA	6900
ACAAGCTCAC TAATAAGATA ATATACATAT TATGGACAAT AATCCTGGTG TTATTATCAA	6960
TAGTCTTCAT CATAGTGCTA ATTAATTCCA TCAAAAGTGA AAAGGCCAC GAATCATTGC	7020
TGCAAGACAT AAATAATGAG TTTATGGAAA TTACAGAAAA GATCCAAATG GCATCGGATA	7080
ATACCAATGA TCTAATACAG TCAGGAGTGA ATACAAGGCT TCTTACAATT CAGAGTCATG	7140
TCCAGAATTA CATAACCAATA TCATTGACAC AACAGATGTC AGATCTTAGG AAATTCAATT	7200
GTGAAATTAC AATTAGAAAAT GATAATCAAG AAGTGCTGCC ACAAAAGAATA ACACATGATG	7260
TAGGTATAAA ACCTTTAAAT CCAGATGATT TTTGGAGATG CACGTCTGGT CTTCCATCTT	7320
TAATGAAAAC TCCAAAATA AGGTTAATGC CAGGGCCGGG ATTATTAGCT ATGCCAACGA	7380
CTGTTGATGG CTGTGTTAGA ACTCCGTCTT TAGTTATAAA TGATCTGATT TATGCTTATA	7440
CCTCAAATCT AATTACTCGA GGTTGTCAGG ATATAGGAAA ATCATATCAA GTCTTACAGA	7500
TAGGGATAAT AACTGTAAAC TCAGACTTGG TACCTGACTT AAATCCTAGG ATCTCTCATA	7560
CCTTTAACAT AAATGACAAT AGGAAGTCAT GTTCTCTAGC ACTCCTAAAT ATAGATGTAT	7620
ATCAACTGTG TTCAACTCCC AAAGTTGATG AAAGATCAGA TTATGCATCA TCAGGCATAG	7680
AAGATATTGT ACTTGATATT GTCAATTATG ATGGTTCAAT CTCACAAACA AGATTTAAGA	7740
ATAATAACAT AAGCTTGAT CAACCATATG CTGCACTATA CCCATCTGTT GGACCAGGGA	7800
TATACTACAA AGGCAAATAA ATATTCTCG GGTATGGAGG TCTTGAACAT CCAATAAAATG	7860
AGAATGTAAT CTGCAACACA ACTGGGTGCC CCGGGAAAAC ACAGAGAGAC TGTAATCAAG	7920
CATCTCATAG TACTGGTTT TCAGATAGGA GGATGGTCAA CTCCATCATT GTTGTGACA	7980
AAGGCTTAAA CTCATTCCA AAATTGAAAG TATGGACGAT ATCTATGCGA CAAAATTACT	8040
GGGGGTCAGA AGGAAGGTTA CTTCTACTAG GTAACAAGAT CTATATATAT ACAAGATCTA	8100
CAAGTTGGCA TAGCAAGTTA CAATTAGGAA TAATTGATAT TACTGATTAC AGTGATATAA	8160

GGATAAAATG GACATGGCAT AATGTGCTAT CAAGACCAGG AAACAATGAA TGTCCATGGG	8220
GACATTCATG TCCAGATGGA TGTATAACAG GAGTATATAC TGATGCATAT CCACTCAATC	8280
CCACAGGGAG CATTGTGTCA TCTGTATAT TAGACTCACA AAAATCGAGA GTGAACCCAG	8340
TCATAACTTA CTCAACAGCA ACCGAAAGAG TAAACGAGCT GGCCATCCTA AACAGAACAC	8400
TCTCAGCTGG ATATACAACA ACAAGCTGCA TTACACACTA TAACAAAGGA TATTGTTTC	8460
ATATAGTAGA AATAAATCAT AAAAGCTTAA ACACATTTCA ACCCATGTTG TTCAAAACAG	8520
AGATTCCAAA AAGCTGCAGT TAATCATAAT TAACCATAAT ATGCATCAAT CTATCTATAA	8580
TACAAGTATA TGATAAGTAA TCAGCAATCA GACAATAGAC AAAAGGGAAA TATAAAAAC	8640
TTAGGAGCAA AGCGTGCTCG GGAAATGGAC ACTGAATCTA ACAATGGCAC TGTATCTGAC	8700
ATACTCTATC CTGAGTGTCA CCTTAACACT CCTATCGTTA AAGGTAAAAT AGCACAATTA	8760
CACACTATTA TGAGTCTACC TCAGCCTTAT GATATGGATG ACGACTCAAT ACTAGTTATC	8820
ACTAGACAGA AAATAAAAAT TAATAAATTG GATAAAAGAC AACGATCTAT TAGAAGATTA	8880
AAATTAATAT TAACTGAAAA AGTGAATGAC TTAGGAAAAT ACACATTTAT CAGATATCCA	8940
GAAATGTCAA AAGAAATGTT CAAATTATAT ATACCTGGTA TTAACAGTAA AGTGAUTGAA	9000
TTATTACTTA AAGCAGATAG AACATATAGT CAAATGACTG ATGGATTAAG AGATCTATGG	9060
ATTAATGTGC TATCAAAATT AGCCTCAAAA AATGATGGAA GCAATTATGA TCTTAATGAA	9120
GAAATTAATA ATATATCGAA AGTTCACACCA ACCTATAAAT CAGATAATG GTATAATCCA	9180
TTCAAAACAT GGTACTAT CAAGTATGAT ATGAGAAGAT TACAAAAAGC TCGAAATGAG	9240
ATCACTTTA ATGTTGGAA GGATTATAAC TTGTTAGAAG ACCAGAAGAA TTTCTTATTG	9300
ATACATCCAG AATTGGTTTT GATATTAGAT AAACAAAATC ATAATGGTTA TCTAATTACT	9360
CCTGAATTAG TATTGATGTA TTGTGACGTA GTCGAAGGCC GATGGAATAT AAGTGCATGT	9420
GCTAAGTTAG ATCCAAAATT ACAATCTATG TATCAGAAAG GTAATAACCT GTGGGAAGTG	9480
ATAGATAAAAT TGTTCCAAT TATGGGAGAA AAGACATTG ATGTGATATC GTTATTAGAA	9540
CCACTTGCAT TATCCTTAAT TCAAACATCAT GATCCTGTTA AACAACTAAG AGGAGCTTT	9600
TTAAATCATG TGTTATCCGA GATGGAATTA ATATTTGAAT CTAGAGAATC GATTAAGGAA	9660
TTTCTGAGTG TAGATTACAT TGATAAAATT TTAGATATAT TTAATAAGTC TACAATAGAT	9720
GAAATAGCAG AGATTTCTC TTTTTTAGA ACATTTGGC ATCCTCCATT AGAAGCTAGT	9780
ATTGCAGCAG AAAAGGTTAG AAAATATATG TATATTGGAA AACAAATTAAA ATTTGACACT	9840
ATTAATAAAAT GTCATGCTAT CTTCTGTACA ATAATAATTA ACGGATATAG AGAGAGGCAT	9900
GGTGGACAGT GCCCTCCTGT GACATTACCT GATCATGCAC ACGAATTCAAT CATAAATGCT	9960
TACGGTTCAA ACTCTGCGAT ATCATATGAA AATGCTGTTG ATTATTACCA GAGCTTTATA	10020
GGAATAAAAT TCAATAAATT CATAGAGCCT CAGTTAGATG AGGATTTGAC AATTTATATG	10080
AAAGATAAAAG CATTATCTCC AAAAAAATCA AATTGGGACA CAGTTTATCC TGCATCTAAT	10140
TTACTGTACC GTACTAACGC ATCCAACGAA TCACGAAGAT TAGTTGAAGT ATTTATAGCA	10200
GATAGTAAAT TTGATCCTCA TCAGATATTG GATTATGTAG AATCTGGGGA CTGGTTAGAT	10260

GATCCAGAAT TTAATATTTT TTATAGTCCTT AAAGAAAAAG AGATCAAACA GGAAGGTAGA	10320
CTCTTGCCTAA AAATGACATA CAAAATGAGA GCTACACAAG TTTTATCAGA GACCCTACTT	10380
GCAAATAACA TAGGAAAATT CTTTCAAGAA AATGGGATGG TGAAGGGAGA GATTGAATTA	10440
CTTAAGAGAT TAACAACCCT ATCAATATCA GGAGTTCCAC GGTATAATGA AGTGTACAAT	10500
AATTCTAAAA GCCATACAGA TGACCTTAAA ACCTACAATA AAATAAGTAA TCTTAATTTG	10560
TCTTCTAATC AGAAATCAA GAAATTGAA TTCAAGTCAA CGGATATCTA CAATGATGGA	10620
TACGAGACTG TGAGCTGTTT CCTAACACA GATCTAAAA AATACTGTCT TAATTGGAGA	10680
TATGAATCAA CAGCTCTATT TGGAGAAACT TGCAACCAAA TATTTGGATT AAATAAATTG	10740
TTTAATTGGT TACACCCTCG TCTTGAAGGA AGTACAATCT ATGTAGGTGA TCCTTACTGT	10800
CCTCCATCAG ATAAAGAACAA TATATCATTAA GAGGATCACC CTGATTCTGG TTTTTACGTT	10860
CATAACCCAA GAGGGGGTAT AGAAGGATTG TGTCAAAAAT TATGGACACT CATATCTATA	10920
AGTGCAATAC ATCTAGCAGC TGTTAGAATA GGCGTGAGGG TGACTGCAAT GGTTCAAGGA	10980
GACAATCAAG CTATAGCTGT AACCACAAGA GTACCCAACA ATTATGACTA CAGAGTTAAG	11040
AAGGAGATAG TTTATAAAGA TGTAGTGAGA TTTTTGATT CATTAAAGAGA AGTGATGGAT	11100
GATCTAGGTC ATGAACCTAA ATTAAATGAA ACGATTATAA GTAGCAAGAT GTTCATATAT	11160
AGCAAAAGAA TCTATTATGA TGGGAGAATT CTTCCCTCAAG CTCTAAAAGC ATTATCTAGA	11220
TGTGTCTTCT GGTCAAGAGAC AGTAATAGAC GAAACAAGAT CAGCATCTTC AAATTTGGCA	11280
ACATCATTG CAAAAGCAAT TGAGAATGGT TATTCACCTG TTCTAGGATA TGCACTGCTCA	11340
ATTTTTAAGA ATATTCAACA ACTATATATT GCCCTGGGA TGAATATCAA TCCAACCTATA	11400
ACACAGAATA TCAGAGATCA GTATTTAGG AATCCAAATT GGATGCAATA TGCCTCTTTA	11460
ATACCTGCTA GTGTTGGGG ATTCAATTAC ATGCCATGT CAAGATGTTT TGTAAGGAAT	11520
ATTGGTGATC CATCAGTTGC CGCATTGGCT GATATTAAA GATTTATTAA GGCGAATCTA	11580
TTAGACCGAA GTGTTCTTTA TAGGATTATG AATCAAGAAC CAGGTGAGTC ATCTTTTTG	11640
GACTGGGCTT CAGATCCATA TTCATGCAAT TTACCCAAAT CTCAAAATAT AACCAACCATG	11700
ATAAAAAAATA TAACAGCAAG GAATGTATTA CAAGATTAC CAAATCCATT ATTATCTGGA	11760
TTATTACCAA ATACAATGAT AGAAGAAGAT GAAGAATTAG CTGAGTTCCCT GATGGACAGG	11820
AAGGTAATTC TCCCTAGAGT TGCACTGAT ATTCTAGATA ATTCTCTCAC AGGAATTAGA	11880
AATGCCATAG CTGGAATGTT AGATACGACA AAATCACTAA TTGGGTTGG CATAAATAGA	11940
GGAGGACTGA CATATAGTTT GTTGAGGAAA ATCAGTAATT ACGATCTAGT ACAATATGAA	12000
ACACTAAGTA GGACTTGGCG ACTAATTGTA AGTGATAAAA TCAAGTATGA AGATATGTGT	12060
TCGGTAGACC TTGCCATAGC ATTGCGACAA AAGATGTGGA TTCATTTATC AGGAGGAAGG	12120
ATGATAAGTG GACTTGAAAC GCCTGACCCA TTAGAATTAC TATCTGGGGT AGTAATAACA	12180
GGATCAGAAC ATTGTAAAAT ATGTTATTCT TCAGATGGCA CAAACCCATA TACTGGATG	12240
TATTTACCCG GTAATATCAA AATAGGATCA GCAGAACAG GTATATCGTC ATTAAGAGTT	12300
CCTTATTTG GATCAGTCAC TGATGAAAGA TCTGAAGCAC AATTAGGATA TATCAAGAAC	12360

CTTAGTAAAC CTGCAAAAGC CGCAATAAGA ATAGCAATGA TATATACATG GGCATTTGGT	12420
AATGATGAGA TATCTTGGAT GGAAGCCTCA CAGATAGCAC AAACACGTGC AAATTTACA	12480
CTAGATAGTC TCAAAATTTC AACACCGGT A GCTACATCAA CAAATTATC ACACAGATTA	12540
AAGGATACTG CAACTCAGAT GAAATTCTCC AGTACATCAT TGATCAGAGT CAGCAGATT	12600
ATAACAATGT CCAATGATAA CATGTCTATC AAAGAACGTA ATGAAACCAA AGATACTAAT	12660
CTTATTATC AACAAATAAT GTTAACAGGA TTAAGTGTTC TCGAATATTT ATTTAGATTA	12720
AAAGAAACCA CAGGACACAA CCCTATAGTT ATGCATCTGC ACATAGAAGA TGAGTGTG	12780
ATTAAAGAAA GTTTAATGA TGAACATATT AATCCAGAGT CTACATTAGA ATTAATTG	12840
TATCCTGAAA GTAATGAATT TATTTATGAT AAAGACCCAC TCAAAGATGT GGACTTATCA	12900
AAACTTATGG TTATTAAAGA CCATTCTTAC ACAATTGATA TGAATTATTG GGATGATACT	12960
GACATCATAAC ATGCAATTTC AATATGTA C GCAATTACAA TAGCAGATAC TATGTCACAA	13020
TTAGATCGAG ATAATTAAA AGAGATAATA GTTATTGCAA ATGATGATGA TATTAATAGC	13080
TTAATCACTG AATTTTGAC TCTTGACATA CTTGTATTTC TCAAGACATT TGGTGGATTA	13140
TTAGTAAATC AATTTGCATA CACTCTTAT AGTCTAAAAA TAGAAGGTAG GGATCTCATT	13200
TGGGATTATA TAATGAGAAC ACTGAGAGAT ACTTCCCATT CAATATTAAA AGTATTATCT	13260
AATGCATTAT CTCATCCTAA AGTATTCAAG AGGTTCTGGG ATTGTGGAGT TTTAAACCC	13320
ATTTATGGTC CTAATACTGC TAGTCAGAC CAGATAAAAC TTGCCCTATC TATATGTGAA	13380
TATTCACTAG ATCTATTATGAGAGATGG TTGAATGGTG TATCACTTGA AATATACATT	13440
TGTGACAGCG ATATGGAAGT TGCAAATGAT AGGAAACAAG CCTTTATTTC TAGACACCTT	13500
TCATTTGTTT GTTGTGTTAGC AGAAATTGCA TCTTCGGAC CTAACCTGTT AAACCTAAC	13560
TACTTGGAGA GACTTGATCT ATTGAAACAA TATCTTGAAT TAAATATTAA AGAAGACCC	13620
ACTCTTAAAT ATGTACAAAT ATCTGGATTA TTAATTAAAT CGTCCCCTC AACTGTAACA	13680
TACGTAAGAA AGACTGCAAT CAAATATCTA AGGATTGCG GTATTAGTCC ACCTGAGGTA	13740
ATTGATGATT GGGATCCGGT AGAAGATGAA AATATGCTGG ATAACATTGT CAAAACATA	13800
AATGATAACT GTAATAAAGA TAATAAAGGG AATAAAATTA ACAATTCTG GGGACTAGCA	13860
CTTAAGAACT ATCAAGTCCT TAAATCAGA TCTATAACAA GTGATTCTGA TGATAATGAT	13920
AGACTAGATG CTAATACAAG TGGTTGACA CTTCTCAAG GAGGGATTAA TCTATCGCAT	13980
CAATTGAGAT TATTGGAAT CAACAGCACT AGTTGCTGA AAGCTCTTGA GTTATCACAA	14040
ATTTTAATGA AGGAAGTCAA TAAAGACAAG GACAGGCTCT TCCTGGGAGA AGGAGCAGGA	14100
GCTATGCTAG CATGTTATGA TGCCACATTA GGACCTGCGAG TTAATTATTA TAATTCAAGGT	14160
TTGAATATAA CAGATGTAAT TGGTCAACGA GAATTGAAAA TATTCCTTC AGAGGTATCA	14220
TTAGTAGGTA AAAAATTAGG AAATGTGACA CAGATTCTTA ACAGGGTAAA AGTACTGTT	14280
AATGGGAATC CTAATTCAAC ATGGATAGGA AATATGGAAT GTGAGAGCTT AATATGGAGT	14340
GAATTAAATG ATAAGTCCAT TGGATTAGTA CATTGTGATA TGGAAAGGAGC TATCGGTAAA	14400
TCAGAAGAAA CTGTTCTACA TGAACATTAT AGTGTATAA GAATTACATA CTTGATTGGG	14460

GATGATGATG TTGTTTACT TTCCAAAATT ATACCTACAA TCACTCCGAA TTGGTCTAGA	14520
ATACTTTATC TATATAAATT ATATTGGAAA GATGTAAGTA TAATATCACT CAAAACCTCT	14580
AATCCTGCAT CAACAGAATT ATATCTAATT TCGAAAGATG CATATTGTAC TATAATGGAA	14640
CCTAGTAAA TTGTTTATC AAAACTTAAA AGATTGTCAC TCTTGGAAAGA AAATAATCTA	14700
TTAAAATGGA TCATTTATC AAAGAAGAGG AATAATGAAT GGTTACATCA TGAAATCAA	14760
GAAGGAGAAA GAGATTATGG AATCATGAGA CCATATCATA TGGCACTACA AATCTTGGA	14820
TTTCAAATCA ATTTAAATCA TCTGGCGAAA GAATTTTAT CAACCCCAGA TCTGACTAAT	14880
ATCAACAATA TAATCCAAAG TTTTCAGCGA ACAATAAAGG ATGTTTATT TGAATGGATT	14940
AATATAACTC ATGATGATAA GAGACATAAA TTAGGCGGAA GATATAACAT ATTCCCAC	15000
AAAATAAGG GAAAGTTAAG ACTGCTATCG AGAAGACTAG TATTAAGTTG GATTTCATTA	15060
TCATTATCGA CTCGATTACT TACAGGTCGC TTTCCTGATG AAAAATTGAA ACATAGAGCA	15120
CAGACTGGAT ATGTATCATT AGCTGATACT GATTTAGAAT CATTAAAGTT ATTGTCGAAA	15180
AACATCATTAGA ATTGATCGGT GGTGCTAAAT TATTAGGAAT TCCCAGACAA	15240
GAAGTTAAAA TACTTATGAA ATTGATCGGT GGTGCTAAAT TATTAGGAAT TCCCAGACAA	15300
TATAAAGAAC CCGAAGACCA GTTATTAGAA AACTACAATC AACATGATGA ATTTGATATC	15360
GATTAAAACA TAAATACAAT GAAGATATAT CCTAACCTTT ATCTTAAGC CTAGGAATAG	15420
ACAAAAAGTA AGAAAAACAT GTAATATATA TATACAAAC AGAGTTCTTC TCTTGTGTTGG	15480
TGGGTGGCA TGGCATCTCC ACCTCCTCGC GGTCCGGACC TGGGCATCCG AAGGAGGACG	15540
CACGTCCACT CGGATGGCTA AGGGAGAGCC TGCAGTAGCA TAACCCCTTG GGGCCTCTAA	15600
ACGGGTCTTG AGGGGTTTT TGCTGAAAGG AGGAACATATA TACCGCGTCGA CGGGCCCCGC	15660
GCTCAC	15666

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GGATTTGCGC GCAATTAAA TCATCTGG

28

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 70 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CCCAGGTCGG ACCCGCAGGGA GGTGGAGATG CCATGCCAGC CCACCAAAAC AAGAGAAGAA	60
CTCTGTTGG	70

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 45 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GGCCCGTCGA CGCGTAATAC GACTCACTAT AGGACCAAAC AAGAG	45
---	----

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 16 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CGGCATCACG TGCTAC	16
-------------------	----

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 6843 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ATGTTGAGCC TATTTGATAC ATTTAATGCA CGTAGGCAAG AAAACATAAC AAAATCAGCC	60
GGTGGAGCTA TCATTCCTGG ACAGAAAAAT ACTGTCTCTA TATTGCCCT TGGACCGACA	120
ATAACTGATG ATAATGAGAA AATGACATTA GCTCTTCTAT TTCTATCTCA TTCACTAGAT	180
AATGAGAAC AACATGCACA AAGGGCAGGG TTCTGGTGT CTTTATTGTC AATGGCTTAT	240
GCCAATCCAG AGCTCTACCT ACAACAAAT GGAAGTAATG CAGATGTCAA GTATGTCATA	300
TACATGATTG AGAAAAGATCT AAAACGGCAA AAGTATGGAG GATTTGTGGT TAAGACGAGA	360
GAGATGATAT ATGAAAAGAC AACTGATTGG ATATTTGGAA GTGACCTGGA TTATGATCAG	420
GAAACTATGT TGCAGAACGG CAGGAACAAT TCAACAATTG AAGACCTTGT CCACACATTT	480

GGGTATCCAT CATGTTAGG AGCTCTTATA ATACAGATCT GGATAGTTCT GGTCAAAGCT	540
ATCACTAGTA TCTCAGGGTT AAGAAAAGGC TTTTCACCC GATTGGAAGC TTTCAGACAA	600
GATGGAACAG TGCAGGCAGG GCTGGTATTG AGCGGTGACA CAGTGGATCA GATTGGGTCA	660
ATCATGCGGT CTCAACAGAG CTTGGTAACT CTTATGGTTG AAACATTAAT AACAAATGAAT	720
ACCAGCAGAA ATGACCTCAC AACCATAGAA AAGAATATAC AAATTGTTGG CAACTACATA	780
AGAGATGCAG GTCTCGCTTC ATTCTTCAAT ACAATCAGAT ATGGAATTGA GACCAGAATG	840
GCAGCTTGA CTCTATCCAC TCTCAGACCA GATATCAATA GATTAAAAGC TTTGATGGAA	900
CTGTATTAT CAAAGGGACC ACGCGCTCCT TTCATCTGTA TCCTCAGAGA TCCTATACAT	960
GGTGAGTTCG CACCAGGCAA CTATCCTGCC ATATGGAGCT ATGCAATGGG GGTGGCAGTT	1020
GTACAAAATA GAGCCATGCA ACAGTATGTG ACGGGAAAGAT CATACTAGA CATTGATATG	1080
TTCCAGCTAG GACAAGCAGT AGCACGTGAT GCCGAAGCTC AAATGAGCTC AACACTGGAA	1140
GATGAACCTG GAGTGACACA CGAACATCTAAA GAAAGCTTGA AGAGACATAT AAGGAACATA	1200
AACAGTTCAAG AGACATCTTT CCACAAACCG ACAGGTGGAT CAGGCATAGA GATGGCAATA	1260
GATGAAGAGC CAGAACAAATT CGAACATAGA GCAGATCAAG AACAAAATGG AGAACCTCAA	1320
TCATCCATAA TTCAATATGC CTGGGCAGAA GGAAATAGAA GCGATGATCA GACTGAGCAA	1380
GCTACAGAAT CTGACAATAT CAAGACCGAA CAACAAAACA TCAGAGACAG ACTAAACAAG	1440
AGACTCAACG ACAAGAAGAA ACAAAAGCAGT CAACCACCC CTAATCCCAC AAACAGAACAA	1500
AACCAGGACG AAATAGATGA TCTGTTAAC GCATTTGGAA GCAACTAAGT CGACGATCCG	1560
GCTGCTAACCA AAGCCCCAAA GGAAGCTGAG TTGGCTGCTG CCACCGCTGA GCAATAACTA	1620
GCATAACCCCC TTGGGGCCTC TAAACGGGTC TTGAGGGGTT TTTGCTGAA AGGAGGAAC	1680
ATATCCGGAT CGAGATCAAT TCTGTGAGCG TATGGCAAAAC GAAGGAAAAAA TAGTTATAGT	1740
AGCCGCACTC GATGGGACAT TTCAACGTAA ACCGTTAAC AATATTTGA ATCTTATTCC	1800
ATTATCTGAA ATGGTGGTAA AACTAACTGC TGTGTGTATG AAATGCTTTA AGGAGGCTTC	1860
CTTTCTAAA CGATTGGGTG AGGAAACCGA GATAGAAATA ATAGGAGGTA ATGATATGTA	1920
TCAATCGGTG TGTAGAAAGT GTTACATCGA CTCATAATAT TATATTTTT ATCTAAAAAA	1980
CTAAAAATAA ACATTGATTA AATTTAATA TAATACTAA AAATGGATGT TGTGTCGTTA	2040
GATAAACCGT TTATGTATT TGAGGAAATT GATAATGAGT TAGATTACGA ACCAGAAAGT	2100
GCAAATGAGG TCGAAAAAA ACTGCCGTAT CAAGGACAGT TAAAACATT ACTAGGAGAA	2160
TTATTTTTC TTAGTAAGTT ACAGCGACAC GGTATATTAG ATGGTGCCAC CGTAGTGTAT	2220
ATAGGATCTG CTCCCGGTAC ACATATACGT TATTTGAGAG ATCATTCTA TAATTTAGGA	2280
GTGATCATCA AATGGATGCT AATTGACGGC CGCCATCATG ATCCTATTTT AAATGGATTG	2340
CGTGATGTGA CTCTAGTGAC TCGGTTCGTT GATGAGGAAT ATCTACGATC CATCAAAAAAA	2400
CAAATGCATC CTTCTAAGAT TATTTAATT TCTGATGTGA GATCCAAACG AGGAGGAAAT	2460
GAACCTAGTA CGGCGGATT ACTAAGTAAT TACGCTCTAC AAAATGTCAT GATTAGTATT	2520
TTAAACCCCG TGGCGTCTAG TCTTAAATGG AGATGCCGT TTCCAGATCA ATGGATCAAG	2580

GACTTTATA TCCCACACGG TAATAAAATG TTACAACCTT TTGCTCCTTC ATATTCAAGG	2640
CCGTCGTTT ACAACGTCGT GACTGGAAA ACCCTGGCGT TACCCAACCTT AATGCCCTG	2700
CAGCACATCC CCCTTCGCC AGCTGGCGTA ATAGCGAAGA GGCCCGCACC GATGCCCTT	2760
CCCAACAGTT GCGCAGCCTG AATGGCGAAT GGCGCGACGC GCCCTGTAGC GGCGCATTAA	2820
GCGCGGGGGG TGTGGTGGTT ACGCGCAGCG TGACCGCTAC ACTTGCCAGC GCCCTAGCGC	2880
CCGCTCCTT CGCTTCTTC CCTTCCTTTC TCGCCACGTT CGCCGGCTTT CCCC GTCAAG	2940
CTCTAAATCG GGGGCTCCCT TTAGGGTTCC GATTTAGTGC TTTACGGCAC CTCGACCCCA	3000
AAAAACTTGA TTAGGGTGAT GGTTCACGTA GTGGGCCATC GCCCTGTAG ACGGTTTTTC	3060
GCCCTTGAC GTTGGAGTCC ACGTTCTTA ATAGTGGACT CTTGTTCCAA ACTGGAACAA	3120
CACTCAACCC TATCTCGGTC TATTCTTTG ATTTATAAGG GATTTGCCG ATTTCGGCCT	3180
ATTGGTTAAA AAATGAGCTG ATTTAACAAA AATTTAACGC GAATTTAAC AAAATATTAA	3240
CGTTTACAAT TTCCCAGGTG GCACTTTCG GGGAAATGTG CGCGGAACCC CTATTTGTTT	3300
ATTTTCTAA ATACATTCAA ATATGTATCC GCTCATGAGA CAATAACCT GATAAATGCT	3360
TCAATAATAT TGAAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTCG CCCTTATTCC	3420
CTTTTTGCG GCATTTGCC TTCCTGTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA	3480
AGATGCTGAA GATCAGTTGG GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG	3540
TAAGATCCTT GAGAGTTTC GCCCCGAAGA ACGTTTCCA ATGATGAGCA CTTTTAAAGT	3600
TCTGCTATGT GGC CGGTAT TATCCGTAT TGACGCCGG CAAGAGCAAC TCGGTGCGCG	3660
CATACACTAT TCTCAGAATG ACTTGGTTGA GTACTCACCA GTCACAGAAA AGCATCTTAC	3720
GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA ACCATGAGTG ATAACACTGC	3780
GGCCAACCTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTGCACAA	3840
CATGGGGGAT CATGTAACTC GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC	3900
AAACGACGAG CGTGACACCA CGATGCCTGT AGCAATGGCA ACAACGTTGC GCAAACATT	3960
AACTGGCGAA CTACTTACTC TAGCTTCCG GCAACAATTA ATAGACTGGA TGGAGGCCGA	4020
TAAAGTTGCA GGACCACTTC TGCGCTCGC CCTTCCGGCT GGCTGGTTA TTGCTGATAA	4080
ATCTGGAGCC GGTGAGCGTG GGTCTCGCG TATCATTGCA GCACTGGGGC CAGATGGTAA	4140
GCCCTCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG GCAACTATGG ATGAACGAAA	4200
TAGACAGATC GCTGAGATAG GTGCCTCACT GATTAAGCAT TGGTAACTGT CAGACCAAGT	4260
TTACTCATAT ATACTTTAGA TTGATTTAAA ACTTCATTAA TAATTTAAAA GGATCTAGGT	4320
GAAGATCCTT TTTGATAATC TCATGACCAA AATCCCTTAA CGTGAGTTT CGTTCCACTG	4380
AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA GATCCTTTT TTCTGCGCGT	4440
AATCTGCTGC TTGCAAACAA AAAAACCAAC GCTACCAGCG GTGGTTGTT TGCCGGATCA	4500
AGAGCTACCA ACTCTTTTC CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAAATAC	4560
TGTCCTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCCTAC	4620
ATACCTCGCT CTGCTAATCC TGTTACCAAGT GGCTGCTGCC AGTGGCGATA AGTCGTGTCT	4680

TACCGGGTTG	GACTCAAGAC	GATAGTTACC	GGATAAGGCG	CAGCGGTCGG	GCTGAACGGG	4740
GGGTTCGTGC	ACACAGCCC	GCTTGAGCG	AACGACCTAC	ACCGAACTGA	GATACCTACA	4800
GCGTGAGCTA	TGAGAAAGCG	CCACGCTTCC	CGAAGGGAGA	AAGGCGGACA	GGTATCCGGT	4860
AAGCGGCAGG	GTCGGAACAG	GAGAGCGCAC	GAGGGAGCTT	CCAGGGGGAA	ACGCCTGGTA	4920
TCTTTATAGT	CCTGTCGGGT	TTCGCCACCT	CTGACTTGAG	CGTCGATT	TGTGATGCTC	4980
GTCAGGGGGG	CGGAGCCTAT	GGAAAAACGC	CAGCAACGCG	GCCTTTTAC	GGTCCTGGC	5040
CTTTTGCTGG	CCTTTGCTC	ACATGTTCTT	TCCTGCGTTA	TCCCCTGATT	CTGTGGATAA	5100
CCGTATTACC	GCCTTGAGT	GAGCTGATAC	CGCTGCCGC	AGCCGAACGA	CCGAGCGCAG	5160
CGAGTCAGTG	AGCGAGGAAG	CGGAAGAGCG	CCCAATACGC	AAACCGCCTC	TCCCCGCGCG	5220
TTGGCCGATT	CATTAATGCA	GCTGGCACGA	CAGGTTCCC	GACTGGAAAG	CGGGCAGTGA	5280
GCGCAACGCA	ATTAATGTGA	GTTAGCTCAC	TCATTAGGCA	CCCCAGGCTT	TACACTTTAT	5340
GCTTCCGGCT	CGTATGTTGT	GTGGAATTGT	GAGCGGATAA	CAATTTCACA	CAGGAAACAG	5400
CTATGACCAT	GATTACGCCA	AGCTTTGCG	ATCAATAAAT	GGATCACAAC	CAGTATCTCT	5460
TAACGATGTT	CTTCGCAGAT	GATGATTCAT	TTTTTAAGTA	TTGGCTAGT	CAAGATGATG	5520
AATCTTCATT	ATCTGATATA	TTGCAAATCA	CTCAATATCT	AGACTTCTG	TTATTATTAT	5580
TGATCCAATC	AAAAAATAAA	TTAGAAGCCG	TGGGTCAATTG	TTATGAATCT	CTTTCAGAGG	5640
AATACAGACA	ATTGACAAAAA	TTCACAGACT	TTCAAGATT	AAAAAAACTG	TTAACAAAGG	5700
TCCCTATTGT	TACAGATGGA	AGGGTCAAAC	TTAATAAAGG	ATATTGTT	GACTTTGTGA	5760
TTAGTTGAT	GCGATTCAA	AAAGAATCCT	CTCTAGCTAC	CACCGCAATA	GATCCTGTTA	5820
GATACATAGA	TCCTCGTCGC	AATATCGCAT	TTTCTAACGT	GATGGATATA	TTAAAGTCGA	5880
ATAAAAGTGA	CAATAATTAA	TTCTTTATTG	TCATCATGAA	CGGGGGACAT	ATTCAAGTTGA	5940
TAATCGGCC	CATGTTTCA	GGTAAAAGTA	CAGAATTAAT	TAGACGAGTT	AGACGTTATC	6000
AAATAGCTCA	ATATAAATGC	GTGACTATAA	AATATTCTAA	CGATAATAGA	TACGGAACGG	6060
GAATATGGAC	GCATGATAAG	AATAATTG	AAGCATTGGA	AGCAACTAAA	CTATGTGATG	6120
TCTTGGAAATC	AATTACAGAT	TTCTCCGTGA	TAGGTATCGA	TGAAGGACAG	TTCTTCCAG	6180
ACATTGTTGA	ATTGATCTCG	ATCCCGCGAA	ATTAATACGA	CTCACTATAG	GGAGACCACA	6240
ACGGTTCCC	TCTAGCGGGA	TCAATTCCGC	CCCTCTCCCT	CCCCCCCCCC	TAACGTTACT	6300
GGCCGAAGCC	GCTTGGAAATA	AGGCCGGTGT	GCGTTGTCT	ATATGTTATT	TTCCACCATA	6360
TTGCCGTCTT	TTGGCAATGT	GAGGGCCCGG	AAACCTGGCC	CTGCTTCTT	GACGAGCATT	6420
CCTAGGGGTC	TTTCCCCCTCT	CGCCAAAGGA	ATGCAAGGTC	TGTTGAATGT	CGTGAAGGAA	6480
GCAGTTCCCTC	TTGAAGCTTC	TTGAAGACAA	ACAACGTCTG	TAGCGACCT	TTGCAGGCAG	6540
CGGAACCCCC	CACCTGGCGA	CAGGTGCCTC	TGCGGCCAAA	AGCCACGTGT	ATAAGATACA	6600
CCTGCAAAGG	CGGCACAACC	CCAGTGCCAC	GTTGTGAGTT	GGATAGTTGT	GGAAAGAGTC	6660
AAATGGCTCT	CCTCAAGCGT	ATTCAACAAG	GGGCTGAAGG	ATGCCAGAA	GGTACCCCAT	6720
TGTATGGGAT	CTGATCTGGG	GCCTCGGTGC	ACATGCTTAA	CATGTGTTA	GTGAGGTTA	6780

AAAAACGTCT AGGCCCGG AACCACGGGG ACGTGGTTT CCTTGAAAA ACACGATAAT	6840
ACC	6843

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7107 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ATGGAAAGCG ACGCTAAAAA CTATCAAATC ATGGATTCTT GGGAAAGAGGA ATCAAGAGAT	60
AAATCAACTA ATATCTCCTC GGCCCTCAAC ATCATTGAAT TCATACTCAG CACCGACCCC	120
CAAGAAGACT TATCGGAAAA CGACACAATC AACACAAGAA CCCAGCAACT CAGTGCCACC	180
ATCTGTCAAC CAGAAATCAA ACCAACAGAA ACAAGTGAGA AAGATAGTGG ATCAACTGAC	240
AAAAATAGAC AGTCCGGGTC ATCACACGAA TGTACAACAG AAGCAAAAGA TAGAAATATT	300
GATCAGGAAA CTGTACAGAG AGGACCTGGG AGAAGAAGCA GCTCAGATAG TAGAGCTGAG	360
ACTGTGGTCT CTGGAGGAAT CCCCAGAAGC ATCACAGATT CTAAAAATGG AACCCAAAAC	420
ACGGAGGATA TTGATCTCAA TGAAATTAGA AAGATGGATA AGGACTCTAT TGAGGGAAA	480
ATGCGACAAT CTGCAAATGT TCCAAGCGAG ATATCAGGAA GTGATGACAT ATTTACAACA	540
GAACAAAGTA GAAACAGTGA TCATGGAAGA AGCCTGGAAT CTATCAGTAC ACCTGATACA	600
AGATCAATAA GTGTTGTTAC TGCTGCAACA CCAGATGATG AAGAAGAAAT ACTAATGAAA	660
AATAGTAGGA CAAAGAAAAG TTCTTCAACA CATCAAGAAG ATGACAAAAG AATTAAAAAA	720
GGGGGAAAAG GGAAAGACTG GTTTAAGAAA TCAAAAGATA CCGACAAACCA GATACCAACA	780
TCAGACTACA GATCCACATC AAAAGGGCAG AAGAAAATCT CAAAGACAAC AACCAACAC	840
ACCGACACAA AGGGGAAAC AGAAATACAG ACAGAACAT CAGAAACACA ATCCTCATCA	900
TGGAATCTCA TCATCGACAA CAACACCGAC CGGAACGAAC AGACAAGCAC AACTCCTCCA	960
ACAACAACCTT CCAGATCAAC TTATACAAAA GAATCGATCC GAACAAACTC TGAATCCAAA	1020
CCCAAGACAC AAAAGACAAA TGGAAAGGAA AGGAAGGATA CAGAAGAGAG CAATCGATTT	1080
ACAGAGAGGG CAATTACTCT ATTGCAGAAT CTTGGTGTAA TTCAATCCAC ATCAAAACTA	1140
GATTTATATC AAGACAAACG AGTTGTATGT GTAGCAAATG TACTAAACAA TGTAGATACT	1200
GCATCAAAGA TAGATTTCTT GGCAGGATTA GTCATAGGGG TTTCAATGGA CAACGACACA	1260
AAATTAACAC AGATACAAAAA TGAAATGCTA AACCTCAAAG CAGATCTAAA GAAAATGGAC	1320
GAATCACATA GAAGATTGAT AGAAAATCAA AGAGAACAAAC TGTCAATTGAT CACGTCACTA	1380
ATTTCAAATC TCAAAATTAT GACTGAGAGA GGAGGAAAGA AAGACCAAAA TGAATCCAAT	1440
GAGAGAGTAT CCATGATCAA AACAAAATTG AAAGAAGAAA AGATCAAGAA GACCAGGTTT	1500
GACCCACTTA TGGAGGCACA AGGCATTGAC AAGAATATAC CCGATCTATA TCGACATGCA	1560

GGAGATAACAC TAGAGAACGA TGTACAAGTT AAATCAGAGA TATTAAGTTC ATACAATGAG	1620
TCAAATGCAA CAAGACTAAT ACCCAAAAAA GTGAGCAGTA CAATGAGATC ACTAGTTGCA	1680
GTCATCAACA ACAGCAATCT CTCACAAAGC ACAAAACAAT CATAACATAAA CGAACTCAAA	1740
CGTTGCAAAA ATGATGAAGA AGTATCTGAA TTAATGGACA TGTTCAATGA AGATGTCAAC	1800
AATTGCCAAT GAGTCGACGA TCCGGCTGCT AACAAAGCCC GAAAGGAAGC TGAGTTGGCT	1860
GCTGCCACCG CTGAGCAATA ACTAGCATAA CCCCTTGGGG CCTCTAAACG GGTCTTGAGG	1920
GGTTTTTGCG TGAAAGGAGG AACTATATCC GGATCGAGAT CAATTCTGTG AGCGTATGGC	1980
AAACGAAGGA AAAATAGTTA TAGTAGCCGC ACTCGATGGG ACATTTCAAC GTAAACCGTT	2040
TAATAATATT TTGAATCTTA TTCCATTATC TGAAATGGTG GTAAAACCAA CTGCTGTGTG	2100
TATGAAATGC TTTAAGGAGG CTTCCCTTTTCA TAAACGATTG GGTGAGGAAA CCGAGATAGA	2160
AATAATAGGA GGTAAATGATA TGTATCAATC GGTGTGTAGA AAGTGTACA TCGACTCATA	2220
ATATTATATT TTTTATCTAA AAAACTAAAA ATAACATTG ATTAAATTAA AATATAATAC	2280
TTAAAAATGG ATGTTGTGTC GTTAGATAAA CCGTTTATGT ATTTTGAGGA AATTGATAAT	2340
GAGTTAGATT ACGAACAGA AAGTGCATGAGGAGG GAGGTCGAA AAAACTGCC GTATCAAGGA	2400
CAGTTAAAAC TATTACTAGG AGAATTATTT TTTCTTAGTA AGTTACAGCG ACACGGTATA	2460
TTAGATGGTG CCACCGTAGT GTATATAGGA TCTGCTCCCG GTACACATAT ACGTTATTG	2520
AGAGATCATT TCTATAATTG AGGAGTGATC ATCAAATGGA TGCTAATTGA CGGCCGCCAT	2580
CATGATCCTA TTTAAATGG ATTGCGTGAT GTGACTCTAG TGACTCGGTT CGTTGATGAG	2640
GAATATCTAC GATCCATCAA AAAACAATG CATCCTCTA AGATTATTT AATTCTGAT	2700
GTGAGATCCA AACGAGGAGG AAATGAACCT AGTACGGCGG ATTACTAAG TAATTACGCT	2760
CTACAAAATG TCATGATTAG TATTTAAAC CCCGTGGCGT CTAGTCTAA ATGGAGATGC	2820
CCGTTCCAG ATCAATGGAT CAAGGACTTT TATATCCCAC ACGGTAATAA AATGTTACAA	2880
CCTTTGCTC CTTCATATTG AGGGCCGTG TTTTACAACG TCGTGACTGG GAAAACCTG	2940
GCGTTACCCA ACTTAATCGC CTTGCAGCAC ATCCCCCTT CGCCAGCTGG CGTAATAGCG	3000
AAGAGGCCCG CACCGATCGC CCTTCCCAAC AGTTGCGCAG CCTGAATGGC GAATGGCGCG	3060
ACGCGCCCTG TAGCGCGCA TTAAGCGCGG CGGGTGTGGT GGTTACGCGC AGCGTGACCG	3120
CTACACTTGC CAGCGCCCTA GCGCCCGCTC CTTTCGCTTT CTTCCCTTCC TTTCTCGCCA	3180
CGTTGCCCG CTTTCCCCGT CAAGCTCTAA ATCGGGGCT CCCTTAGGG TTCCGATTTA	3240
GTGCTTACG GCACCTCGAC CCCAAAAAC TTGATTAGGG TGATGGTTCA CGTAGTGGC	3300
CATGCCCTG ATAGACGGTT TTTCGCCCTT TGACGTTGGA GTCCACGTTTCTTAAATAGTG	3360
GACTCTGTT CCAAACACTGGA ACAACACTCA ACCCTATCTC GGTCTATTCT TTTGATTAT	3420
AAGGGATTTT GCCGATTCG GCCTATTGGT TAAAAAATGA GCTGATTTAA CAAAAATTAA	3480
ACGCGAATTG TAACAAAATA TTAACGTTA CAATTCCCA GGTGGCACTT TTCGGGGAAA	3540
TGTGCGCGGA ACCCCTATTG TTGATTGTTT CTAAATACAT TCAAATATGT ATCCGCTCAT	3600
GAGACAATAA CCCTGATAAA TGCTTCAATA ATATTGAAAA AGGAAGAGTA TGAGTATTCA	3660

ACATTTCCGT GTCGCCCTTA TTCCCTTTT TCGGGCATTT TGCCCTCCTG TTTTGCTCA	3720
CCCAGAAACG CTGGTGAAAG TAAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGGTTA	3780
CATCGAACTG GATCTCAACA GCGGTAAGAT CCTTGAGAGT TTTGCCCG AAGAACGTTT	3840
TCCAATGATG AGCACTTTA AAGTTCTGCT ATGTGGCGCG GTATTATCCC GTATTGACGC	3900
CGGGCAAGAG CAACTCGGTC GCCGCATAACA CTATTCTCAG AATGACTTGG TTGAGTACTC	3960
ACCAGTCACA GAAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC	4020
CATAACCATG AGTGATAACA CTGCGGCCAA CTTACTTCTG ACAACGATCG GAGGACCGAA	4080
GGAGCTAACCG GCTTTTTGC ACAACATGGG GGATCATGTA ACTCGCCTTG ATCGTTGGGA	4140
ACCGGAGCTG AATGAAGCCA TACCAAACGA CGAGCGTGAC ACCACGATGC CTGTAGCAAT	4200
GGCAACAAACG TTGCGCAAAC TATTAACTGG CGAACTACTT ACTCTAGCTT CCCGGCAACA	4260
ATTAATAGAC TGGATGGAGG CGGATAAAAGT TGCAGGACCA CTTCTGCGCT CGGCCCTTCC	4320
GGCTGGCTGG TTTATTGCTG ATAAATCTGG AGCCGGTGAG CGTGGGTCTC GCGGTATCAT	4380
TGCAGCACTG GGGCCAGATG GTAAGCCCTC CCGTATCGTA GTTATCTACA CGACGGGGAG	4440
TCAGGCAACT ATGGATGAAC GAAATAGACA GATCGCTGAG ATAGGTGCCT CACTGATTAA	4500
GCATTGGTAA CTGTCAGACC AAGTTACTC ATATATACTT TAGATTGATT TAAAACTTCA	4560
TTTTTAATTT AAAAGGATCT AGGTGAAGAT CCTTTTGAT AATCTCATGA CCAAAATCCC	4620
TTAACGTGAG TTTTCGTTCC ACTGAGCGTC AGACCCCGTA GAAAAGATCA AAGGATCTTC	4680
TTGAGATCCT TTTTTCTGC GCGTAATCTG CTGCTTGCAA ACAAAAAAAC CACCGCTTAC	4740
AGCGGTGGTT TGTGCGCG ATCAAGAGCT ACCAACTCTT TTTCCGAAGG TAACTGGCTT	4800
CAGCAGAGCG CAGATACCAA ATACTGTCCT TCTAGTGTAG CCGTAGTTAG GCCACCACCT	4860
CAAGAACTCT GTAGCACCGC CTACATACCT CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC	4920
TGCCAGTGGC GATAAGTCGT GTCTTACCGG GTTGGACTCA AGACGATAGT TACCGGATAA	4980
GGCGCAGCGG TCGGGCTGAA CGGGGGGTTG GTGCACACAG CCCAGCTTGG AGCGAACGAC	5040
CTACACCGAA CTGAGATACC TACAGCGTGA GCTATGAGAA AGCGCCACGC TTCCCGAAGG	5100
GAGAAAGGCG GACAGGTATC CGGTAAGCGG CAGGGTCGGA ACAGGAGAGC GCACGAGGGA	5160
GCTTCCAGGG GGAAACGCCT GGTATCTTTA TAGCCTGTC GGGTTCGCC ACCTCTGACT	5220
TGAGCGTCGA TTTTTGTGAT GCTCGTCAGG GGGGCGGAGC CTATGGAAA ACGCCAGCAA	5280
CGCGGCCTT TTACGGTTCC TGGCCTTTG CTGGCCTTT GCTCACATGT TCTTCTGTC	5340
GTTATCCCCT GATTCTGTGG ATAACCGTAT TACCGCCTTT GAGTGAGCTG ATACCGCTCG	5400
CCGCAGCCGA ACGACCGAGC GCAGCGAGTC AGTGAGCGAG GAAGCGGAAG AGCGCCAAT	5460
ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCAATTAA TGCAGCTGGC ACGACAGGTT	5520
TCCCAGTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC TCACTCATTAA	5580
GGCACCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA TTGTGAGCGG	5640
ATAACAATTT CACACAGGAA ACAGCTATGA CCATGATTAC GCCAAGCTTT TGCGATCAAT	5700
AAATGGATCA CAACCAGTAT CTCTAACGA TGTTCTTCGC AGATGATGAT TCATTTTTA	5760

AGTATTTGGC TAGTCAAGAT GATGAATCTT CATTATCTGA TATATTGCAA ATCACTCAAT	5820
ATCTAGACTT TCTGTTATTA TTATTGATCC AATCAAAAAA TAAATTAGAA GCCGTGGTC	5880
ATTGTTATGA ATCTCTTCAG GAGGAATACA GACAATTGAC AAAATTACAA GACTTTCAAG	5940
ATTTAAAAAA ACTGTTAAC AAGGTCCCTA TTGTTACAGA TGGAAGGGTC AAACCTTAATA	6000
AAGGATATTT GTTCGACTTT GTGATTAGTT TGATGCGATT CAAAAAAGAA TCCTCTCTAG	6060
CTACCACCGC AATAGATCCT GTTAGATACA TAGATCCTCG TCGCAATATC GCATTTCTA	6120
ACGTGATGGA TATATTAAAG TCGAATAAAG TGAACAATAA TTAATTCTTT ATTGTCACTA	6180
TGAACGGCGG ACATATTCAAG TTGATAATCG GCCCCATGTT TTCAGGTAAA AGTACAGAAT	6240
TAATTAGACG AGTTAGACGT TATCAAATAG CTCAATATAA ATGCGTGACT ATAAAATATT	6300
CTAACGATAA TAGATACGGA ACGGGACTAT GGACGCATGA TAAGAATAAT TTTGAAGCAT	6360
TGGAAGCAAC TAAACTATGT GATGTCTTGG AATCAATTAC AGATTTCTCC GTGATAGGTA	6420
TCGATGAAGG ACAGTTCTTT CCAGACATTG TTGAATTGAT CTCGATCCCC CGAAATTAAT	6480
ACGACTCACT ATAGGGAGAC CACAACGGTT TCCCTCTAGC GGGATCAATT CCGCCCCCTCT	6540
CCCTCCCCC CCCCTAACGT TACTGGCCGA AGCCGCTTGG AATAAGGCCG GTGTGCGTTT	6600
GTCTATATGT TATTTTCCAC CATATTGCCG TCTTTGGCA ATGTGAGGGC CCGGAAACCT	6660
GGCCCTGTCT TCTTGACGAG CATTCTAGG GGTCTTCCC CTCTGCCAA AGGAATGCCA	6720
GGTCTGTTGA ATGTCGTGAA GGAAGCAGTT CCTCTGGAAAG CTTCTGAAG ACAAAACAACG	6780
TCTGTAGCGA CCCTTGCAG GCAGCGGAAC CCCCCCACCTG GCGACAGGTG CCTCTGCCGC	6840
CAAAGCCAC GTGTATAAGA TACACCTGCA AAGGCGGCAC AACCCCAAGTG CCACGTTGTG	6900
AGTTGGATAG TTGTGGAAAG AGTCAAATGG CTCTCCTCAA GCGTATTCAA CAAGGGCTG	6960
AAGGATGCCA AGAAGGTACC CCATTGTATG GGATCTGATC TGGGGCCTCG GTGCACATGC	7020
TTTACATGTG TTTAGTCGAG GTTAAAAAAC GTCTAGGCC CCGAACAC GGGGACGTGG	7080
TTTCCTTTG AAAAACACGA TAATACC	7107

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12011 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ATGGACACTG AATCTAACAA TGGCACTGTA TCTGACATAC TCTATCCTGA GTGTCACCTT	60
AACTCTCCTA TCGTTAAAGG TAAAATAGCA CAATTACACA CTATTATGAG TCTACCTCAG	120
CCTTATGATA TGGATGACGA CTCAATACTA GTTATCACTA GACAGAAAAT AAAACTTAAT	180
AAATTGGATA AAAGACAACG ATCTATTAGA AGATTAAT TAATATTAAC TGAAAAAGTG	240
AATGACTTAG GAAAATACAC ATTTATCAGA TATCCAGAAA TGTCAAAAGA AATGTTCAAA	300

TTATATATACTTGGTATTAA CAGTAAAGTG ACTGAATTAT TACTTAAAGC AGATAGAAC	360
TATAGTCAAA TGACTGATGG ATTAAGAGAT CTATGGATTA ATGTGCTATC AAAATTAGCC	420
TCAAAAAATG ATGGAAGCAA TTATGATCTT AATGAAGAAA TTAATAATAT ATCGAAAGTT	480
CACACAAACCT ATAAATCAGA TAAATGGTAT AATCCATTCA AAACATGGTT TACTATCAAG	540
TATGATATGA GAAGATTACA AAAAGCTCGA AATGAGATCA CTTTTAATGT TGGGAAGGAT	600
TATAACTTGT TAGAAGACCA GAAGAATTTC TTATTGATAC ATCCAGAATT GGTTTGATA	660
TTAGATAAAC AAAACTATAA TGGTTATCTA ATTACTCCTG AATTAGTATT GATGTATTGT	720
GACGTAGTCG AAGGCCGATG GAATATAAGT GCATGTGCTA AGTTAGATCC AAAATTACAA	780
TCTATGTATC AGAAAGGTAA TAACCTGTGG GAAGTGATAG ATAAATTGTT TCCAATTATG	840
GGAGAAAAGA CATTGATGT GATATCGTTA TTAGAACAC TTGCATTATC CTTAATTCAA	900
ACTCATGATC CTGTTAAACA ACTAAGAGGA GCTTTTTAA ATCATGTGTT ATCCGAGATG	960
GAATTAATAT TTGAATCTAG AGAACATCGATT AAGGAATTTC TGAGTGTAGA TTACATTGAT	1020
AAAATTCTAG ATATATTAA TAAGTCTACA ATAGATGAAA TAGCAGAGAT TTTCTCTTT	1080
TTTAGAACAT TTGGGCATCC TCCATTAGAA GCTAGTATTG CAGCAGAAAA GGTTAGAAAA	1140
TATATGTATA TTGGAAAACA ATTAAAATTG GACACTATTA ATAAATGTCA TGCTATCTTC	1200
TGTACAATAA TAATTAACGG ATATAGAGAG AGGCATGGTG GACAGTGGCC TCCTGTGACA	1260
TTACCTGATC ATGCACACGA ATTCACTCATA AATGCTTACG GTTCAAACCTC TGCGATATCA	1320
TATGAAAATG CTGTTGATTA TTACCAAGGC TTTATAGGAA TAAAATTCAA TAAATTCTATA	1380
GAGCCTCAGT TAGATGAGGA TTTGACAATT TATATGAAAG ATAAAGCATT ATCTCCAAAA	1440
AAATCAAATT GGGACACAGT TTATCCTGCA TCTAATTAC TGTACCGTAC TAACGCATCC	1500
AACGAATCAC GAAGATTAGT TGAAGTATTG ATAGCAGATA GTAAATTGTA TCCTCATCAG	1560
ATATTGGATT ATGTAGAACATC TGGGGACTGG TTAGATGATC CAGAATTAA TATTCTTAT	1620
AGTCTTAAAG AAAAGAGAT CAAACAGGAA GGTAGACTCT TTGCAAAAT GACATACAAA	1680
ATGAGAGCTA CACAAGTTT ATCAGAGACA CTACTGCAA ATAACATAGG AAAATTCTTT	1740
CAAGAAAATG GGATGGTGAA GGGAGAGATT GAATTACTTA AGAGATTAAC AACCATATCA	1800
ATATCAGGAG TTCCACGGTA TAATGAAGTG TACAATAATT CTAAAAGCCA TACAGATGAC	1860
CTTAAACCT ACAATAAAAT AAGTAATCTT AATTGTCTT CTAATCAGAA ATCAAAGAAA	1920
TTTGAATTCA AGTCAACGGA TATCTACAAT GATGGATACG AGACTGTGAG CTGTTCTTA	1980
ACAACAGATC TCAAAAAATA CTGTCTTAAT TGGAGATATG AATCAACAGC TCTATTGGA	2040
GAAACTTGCA ACCAAATATT TGGATTAAAT AAATTGTTA ATTGGTTACA CCCTCGTCTT	2100
GAAGGAAGTA CAATCTATGT AGGTGATCCT TACTGTCTC CATCAGATAA AGAACATATA	2160
TCATTAGAGG ATCACCCCTGA TTCTGGTTTT TACGTTCTA ACCCAAGAGG GGGTATAGAA	2220
GGATTTGTC AAAAATTATG GACACTCATA TCTATAAGTG CAATACATCT AGCAGCTGTT	2280
AGAATAGGCG TGAGGGTGAC TGCAATGGTT CAAGGAGACA ATCAAGCTAT AGCTGTAACC	2340
ACAAGAGTAC CCAACAAATTA TGACTACAGA GTTAAGAAGG AGATAGTTA TAAAGATGTA	2400

GTGAGATTTT TTGATTCACTT AAGAGAAGTG ATGGATGATC TAGGTCATGA ACTTAAATTA	2460
AATGAAACGA TTATAAGTAG CAAGATGTTG ATATATAGCA AAAGAACATCTA TTATGATGGG	2520
AGAATTCTTC CTCAAGCTCT AAAAGCATTA TCTAGATGTG TCTTCTGGTC AGAGACAGTA	2580
ATAGACGAAA CAAGATCAGC ATCTCAAAT TTGGCAACAT CATTGCAAA AGCAATTGAG	2640
AATGGTTATT CACCTGTTCT AGGATATGCA TGCTCAATT TTAAGAACAT TCAACAACTA	2700
TATATTGCCCT TTGGGATGAA TATCAATCCA ACTATAACAC AGAATATCAG AGATCAGTAT	2760
TTTAGGAATC CAAATTGGAT GCAATATGCC TCTTTAATAC CTGCTAGTGT TGGGGGATTC	2820
AATTACATGG CCATGTCAAG ATGTTTGTA AGGAATATTG GTGATCCATC AGTTGCCGCA	2880
TTGGCTGATA TTAAAAGATT TATTAAGGCG AATCTATTAG ACCGAAGTGT TCTTTATAGG	2940
ATTATGAATC AAGAACCAAG TGAGTCATCT TTTTGGAAC GGGCTTCAGA TCCATATTCA	3000
TGCAATTAC CACAATCTCA AAATATAACC ACCATGATAA AAAATATAAC AGCAAGGAAT	3060
GTATTACAAG ATTCAACAAA TCCATTATTA TCTGGATTAT TCACAAATAC AATGATAGAA	3120
GAAGATGAAG AATTAGCTGA GTTCCTGATG GACAGGAAGG TAATTCTCCC TAGAGTGCA	3180
CATGATATTC TAGATAATTC TCTCACAGGA ATTAGAAATG CCATAGCTGG AATGTTAGAT	3240
ACGACAAAAT CACTAATTGG GGTTGGCATA AATAGAGGAG GACTGACATA TAGTTGTTG	3300
AGGAAAATCA GTAATTACGA TCTAGTACAA TATGAAACAC TAAGTAGGAC TTTGCGACTA	3360
ATTGTAAGTG ATAAAATCAA GTATGAAGAT ATGTGTTGG TAGACCTTGC CATAGCATTG	3420
CGACAAAAGA TGTGGATTCA TTTATCAGGA GGAAGGATGA TAAGTGGACT TGAAACGCCT	3480
GACCCATTAG AATTACTATC TGGGGTAGTA ATAACAGGAT CAGAACATTG TAAAATATGT	3540
TATTCTTCAG ATGGCACAAA CCCATATACT TGGATGTATT TACCCGGTAA TATCAAAATA	3600
GGATCAGCAG AACACAGGTAT ATCGTCATTA AGAGTCCCTT ATTTGGATC AGTCACTGAT	3660
GAAAGATCTG AAGCACAATT AGGATATATC AAGAACCTTA GTAAACCTGC AAAAGCCGCA	3720
ATAAGAACATAG CAATGATATA TACATGGCA TTTGGTAATG ATGAGATATC TTGGATGGAA	3780
GCCTCACAGA TAGCACAAAC ACGTGCAAAT TTTACACTAG ATAGTCTCAA AATTTAACAA	3840
CCGGTAGCTA CATCAACAAA TTTATCACAC AGATTAAGG ATACTGCAAC TCAGATGAAA	3900
TTCTCCAGTA CATCATTGAT CAGAGTCAGC AGATTCTAA CAATGTCCTA TGATAACATG	3960
TCTATCAAAG AAGCTAATGA AACCAAAGAT ACTAATCTTA TTTATCAACA AATAATGTTA	4020
ACAGGATTAA GTGTTTTCGA ATATTATTT AGATTAAGG AAACCACAGG ACACAAACCT	4080
ATAGTTATGC ATCTGCACAT AGAAGATGAG TGTTGTATTA AAGAAAGTTT TAATGATGAA	4140
CATATTAATC CAGAGTCTAC ATTAGAATTA ATTGATATC CTGAAAGTAA TGAATTATTT	4200
TATGATAAAAG ACCCACTCAA AGATGTGGAC TTATCAAAC TTATGGTTAT TAAAGACCAT	4260
TCTTACACAA TTGATATGAA TTATTGGAT GATACTGACA TCATACATGC AATTTCAATA	4320
TGTACTGCAA TTACAATAGC AGATACTATG TCACAATTAG ATCGAGATAA TTTAAAAGAG	4380
ATAATAGTTA TTGCAAATGA TGATGATATT AATAGCTAA TCACGAAATT TTTGACTCTT	4440
GACATACTTG TATTCTCAA GACATTGGT GGATTATTAG TAAATCAATT TGCATACACT	4500

CTTTATAGTC TAAAAATAGA AGGTAGGGAT CTCATTGGG ATTATATAAT GAGAACACTG	4560
AGAGATACTT CCCATTCAAT ATTAAAAGTA TTATCTAATG CATTATCTCA TCCTAAAGTA	4620
TTCAAGAGGT TCTGGGATTG TGGAGTTTA AACCTATTT ATGGTCTAA TACTGCTAGT	4680
CAAGACCAGA TAAAACCTGC CCTATCTATA TGTGAATATT CACTAGATCT ATTTATGAGA	4740
GAATGGTTGA ATGGTGTATC ACTTGAAATA TACATTGTG ACAGCGATAT GGAAGTTGCA	4800
AATGATAGGA ACAAAGCCTT TATTTCTAGA CACCTTCAT TTGTTGTTG TTTAGCAGAA	4860
ATTGCATCTT TCGGACCTAA CCTGTTAAC TTAACATACT TGGAGAGACT TGATCTATTG	4920
AAACAATATC TTGAATTAAA TATTAAAGAA GACCCTACTC TTAAATATGT ACAAAATATCT	4980
GGATTATTAA TAAATCGTT CCCATCAACT GTAACATACG TAAGAAAGAC TGCAATCAA	5040
TATCTAAGGA TTCGCGGTAT TAGTCCACCT GAGGTAATTG ATGATTGGGA TCCGGTAGAA	5100
GATGAAAATA TGCTGGATAA CATTGTAAA ACTATAATG ATAACGTAA TAAAGATAAT	5160
AAAGGGAATA AAATTAACAA TTTCTGGGA CTAGCACTTA AGAACTATCA AGTCCTAAA	5220
ATCAGATCTA TAACAAGTGA TTCTGATGAT AATGATAGAC TAGATGCTAA TACAAGTGGT	5280
TTGACACTTC CTCAGGAGG GAATTATCTA TCGCATCAAT TGAGATTATT CGGAATCAAC	5340
AGCACTAGTT GTCTGAAAGC TCTTGAGTTA TCACAAATTG TAATGAAGGA AGTCAATAAA	5400
GACAAGGACA GGCTCTTCCT GGGAGAAGGA GCAGGAGCTA TGCTAGCATG TTATGATGCC	5460
ACATTAGGAC CTGCAGTTAA TTATTATAAT TCAGGTTGA ATATAACAGA TGTAATTGGT	5520
CAACGAGAAT TGAAAATATT TCCTTCAGAG GTATCATTAG TAGGTAAAAA ATTAGGAAAT	5580
GTGACACAGA TTCTTAACAG GGTAAGATA CTGTTCAATG GGAATCCTAA TTCAACATGG	5640
ATAGGAAATA TGGAATGTGA GAGCTTAATA TGGAGTGAAT TAAATGATAA GTCCATTGGA	5700
TTAGTACATT GTGATATGGA AGGAGCTATC GGTAAATCAG AAGAAACTGT TCTACATGAA	5760
CATTATAGTG TTATAAGAAT TACATACTTG ATTGGGGATG ATGATGTTGT TTTAGTTCC	5820
AAAATTATAC CTACAATCAC TCCGAATTGG TCTAGAATAC TTTATCTATA TAAATTATAT	5880
TGGAAAGATG TAAGTATAAT ATCACTAAA ACTTCTAATC CTGCATCAAC AGAATTATAT	5940
CTAATTTCGA AAGATGCATA TTGTACTATA ATGGAACCTA GTGAAATTGT TTTATCAAAA	6000
CTTAAAAGAT TGTCACTCTT GGAAGAAAAT AATCTATTAA AATGGATCAT TTTATCAAAG	6060
AAGAGGAATA ATGAATGGTT ACATCATGAA ATCAAAGAAG GAGAAAGAGA TTATGGAATC	6120
ATGAGACCAT ATCATATGGC ACTACAAATC TTTGGATTTC AAATCAATT AAATCATCTG	6180
GCGAAAGAAT TTTTATCAAC CCCAGATCTG ACTAATATCA ACAATATAAT CCAAAGTTT	6240
CAGCGAACAA TAAAGGATGT TTTATTGAA TGGATTAAATA TAACTCATGA TGATAAGAGA	6300
CATAAAATTAG GCGGAAGATA TAACATATTC CCACTGAAAA ATAAGGGAAA GTTAAGACTG	6360
CTATCGAGAA GACTAGTATT AAGTTGGATT TCATTATCAT TATCGACTCG ATTACTTACA	6420
GGTCGCTTTC CTGATGAAAA ATTTGAACAT AGAGCACAGA CTGGATATGT ATCATTAGCT	6480
GATACTGATT TAGAATCATT AAAGTTATTG TCGAAAAACA TCATTAAGAA TTACAGAGAG	6540
TGTATAGGAT CAATATCATA TTGGTTCTA ACCAAAGAAG TTAAAATACT TATGAAATTG	6600

ATTGGTGGTG	CTAAATTATT	AGGAATTCCC	AGACAATATA	AAGAACCGA	AGACCAGTTA	6660
TTAGAAA	ACT ACAATCAACA	TGATGAATT	T GATATCGATT	AAAACATAAA	TACAATGTCG	6720
ACGATCCGGC	TGCTAACAAA	GCCCCGAAAGG	AAGCTGAGTT	GGCTGCTGCC	ACCGCTGAGC	6780
AATAACTAGC	ATAACCCCTT	GGGGCCTCTA	AACGGGTCTT	GAGGGGTTTT	TTGCTGAAAG	6840
GAGGA	ACTAT ATCCGGATCG	AGATCAATT	TGTGAGCGTA	TGGCAAACGA	AGGAAAAATA	6900
GTTATAGTAG	CCGCACTCGA	TGGGACATT	CAACGTAAAC	CGTTAATAA	TATTTGAAT	6960
CTTATTCCAT	TATCTGAAAT	GGTGGTAAAA	CTAACTGCTG	TGTGTATGAA	ATGCTTTAAG	7020
GAGGCTTCCT	TTTCTAAACG	ATTGGGTGAG	GAAACCGAGA	TAGAAATAAT	AGGAGGTAAT	7080
GATATGTATC	AATCGGTGTG	TAGAAAGTGT	TACATCGACT	CATAATATTA	TATTTTTTAT	7140
CTAAAAAA	ACT AAAAATAAAC	ATTGATTAAA	TTTTAATATA	ATACTTAAAA	ATGGATGTTG	7200
TGTCGTTAGA	TAAACCGTTT	ATGTATTTG	AGGAAATTGA	TAATGAGTTA	GATTACGAAC	7260
CAGAAAGTGC	AAATGAGGTC	GCAAAAAAAC	TGCCGTATCA	AGGACAGTTA	AAACTATTAC	7320
TAGGAGAATT	ATTTTTCTT	AGTAAGTTAC	AGCGACACGG	TATATTAGAT	GGTGCCACCG	7380
TAGTGTATAT	AGGATCTGCT	CCC GG TACAC	ATATACGTTA	TTTGAGAGAT	CATTCTATA	7440
ATTTAGGAGT	GATCATCAAA	TGGATGCTAA	TTGACGGCCG	CCATCATGAT	CCTATTTAA	7500
ATGGATTGCG	TGATGTGACT	CTAGTGACTC	GGTTCGTTGA	TGAGGAATAT	CTACGATCCA	7560
TCAAAAAACA	ACTGCATCCT	TCTAAGATTA	TTTTAATTTC	TGATGTGAGA	TCCAAACGAG	7620
GAGGA	AAATGA ACCTAGTACG	GC GG GATTAC	TAAGTAATTA	CGCTCTACAA	AATGTCATGA	7680
TTAGTATTTT	AAACCCCGTG	GCGTCTAGTC	TTAAATGGAG	ATGCCCGTTT	CCAGATCAAT	7740
GGATCAAGGA	CTTTTATATC	CCACACGGTA	ATAAAATGTT	ACAACCTTT	GCTCCTTCAT	7800
ATTCAGGGCC	GTCGTTTAC	AACGTCGTGA	CTGGGAAAC	CCTGGCGTTA	CCCAACTTAA	7860
TCGCCTGCA	GCACATCCCC	CTT CGCCAG	CTGGCGTAAT	AGCGAAGAGG	CCCGCACCGA	7920
TCGCCCTTCC	CAACAGTTGC	GCAGCCTGAA	TGGCGAATGG	CGCGACGCGC	CCTGTAGCGG	7980
CGCATTAA	AGC GC GGGGTG	TGGTGGTTAC	GCGCAGCGTG	ACCGCTACAC	TTGCCAGCGC	8040
CCTAGCGCCC	GCTCCTTCG	CTTTCTTCCC	TTCTTTCTC	GCCACGTTCG	CCGGCTTTCC	8100
CCGTCAAGCT	CTAAATCGGG	GGCTCCCTT	AGGGTTCCGA	TTTAGTGCTT	TACGGCACCT	8160
CGACCCCCAA	AAACTTGATT	AGGGTGATGG	TTCACGTAGT	GGGCCATCGC	CCTGATAGAC	8220
GGTTTT	TCG	CTT GACGT	TGGAGTCCAC	GTTCTTAAT	AGTGGACTCT	8280
TGGAACAACA	CTCAACCCTA	TCTCGGTCTA	TTCTTTGAT	TTATAAGGGA	TTTGCCGAT	8340
TTCGGCCTAT	TGGTTAAAAA	ATGAGCTGAT	TTAACAAAAA	TTAACGCGA	ATTTAACAA	8400
AATATTAACG	TTTACAATT	CCCAGGTGGC	ACTTTCCGGG	GAAATGTGCG	CGGAACCCCT	8460
ATTTGTTTAT	TTTCTAAAT	ACATTCAAAT	ATGTATCCGC	TCATGAGACA	ATAACCCTGA	8520
TAAATGCTTC	AATAATATTG	AAAAGGAAG	AGTATGAGTA	TTCAACATT	CCGTGTCGCC	8580
CTTATTCCCT	TTTTGCGGC	ATTTGCCTT	CCTGTTTTG	CTCACCCAGA	AACGCTGGTG	8640
AAAGTAAAAG	ATGCTGAAGA	TCAGTTGGGT	GCACGAGTGG	GTTACATCGA	ACTGGATCTC	8700

AACAGCGGTA AGATCCTTGA GAGTTTCGC CCCGAAGAAC GTTTCCAAT GATGAGCACT	8760
TTTAAAGTTC TGCTATGTGG CGCGGTATT A TCCC GTATTG ACGCCGGCA AGAGCAACTC	8820
GGTCGCCGCA TACACTATT TCAGAATGAC TTGGTTGAGT ACTCACCAAGT CACAGAAAAG	8880
CATCTTACGG ATGGCATGAC AGTAAGAGAA TTATGCAGTG CTGCCATAAC CATGAGTGAT	8940
AACACTGCGG CCAACTTACT TCTGACAACG ATCGGAGGAC CGAAGGAGCT AACCGCTTT	9000
TTGCACAACA TGGGGGATCA TGTAACTCGC CTTGATCGTT GGGAACCGGA GCTGAATGAA	9060
GCCATACCAA ACGACGAGCG TGACACCACG ATGCCTGTAG CAATGGCAAC AACGTTGCGC	9120
AAACTATTAA CTGGCGAACT ACTTACTCTA GCTTCCCGGC AACAAATTAAAT AGACTGGATG	9180
GAGGGCGATA AAGTTGCAGG ACCACTTCTG CGCTCGGCCC TTCCGGCTGG CTGGTTTATT	9240
GCTGATAAAAT CTGGAGCCGG TGAGCGTGGG TCTCGCGTA TCATTGCAGC ACTGGGGCCA	9300
GATGGTAAGC CCTCCCGTAT CGTAGTTATC TACACGACGG GGAGTCAGGC AACTATGGAT	9360
GAACGAAATA GACAGATCGC TGAGATAGGT GCCTCACTGA TTAAGCATTG GTAACGTCA	9420
GACCAAGTTT ACTCATATAT ACTTTAGATT GATTAAAAC TTCATTTTA ATTTAAAAGG	9480
ATCTAGGTGA AGATCCTTT TGATAATCTC ATGACCAAAA TCCCTTAACG TGAGTTTCG	9540
TTCCACTGAG CGTCAGACCC CGTAGAAAAG ATCAAAGGAT CTTCTTGAGA TCCTTTTTT	9600
CTGCGCGTAA TCTGCTGCTT GCAAACAAAA AAACCACCGC TACCAGCGGT GGTTTGTG	9660
CCGGATCAAG AGCTACCAAC TCTTTTCCG AAGGTAACTG GCTTCAGCAG AGCGCAGATA	9720
CCAAATACTG TCCTCTAGT GTAGCCGTAG TTAGGCCACC ACTTCAAGAA CTCTGTAGCA	9780
CCGCCTACAT ACCTCGCTCT GCTAACCTG TTACCAAGTGG CTGCTGCCAG TGGCGATAAG	9840
TCGTGTCTTA CCGGGTTGGA CTCAAGACGA TAGTTACCGG ATAAGGCGCA GCGGTGGGC	9900
TGAACGGGG GTTCGTGCAC ACAGCCCAGC TTGGAGCGAA CGACCTACAC CGAACTGAGA	9960
TACCTACAGC GTGAGCTATG AGAAAGCGCC ACGCTTCCCG AAGGGAGAAA GGCGGACAGG	10020
TATCCGGTAA GCGGCAGGGT CGGAACAGGA GAGCGCACGA GGGAGCTTCC AGGGGGAAAC	10080
GCCTGGTATC TTTATAGTCC TGTCGGTTT CGCCACCTCT GACTTGAGCG TCGATTTTG	10140
TGATGCTCGT CAGGGGGCG GAGCCTATGG AAAAACGCCA GCAACGCGGC CTTTTACGG	10200
TTCCTGGCCT TTTGCTGGCC TTTGCTCAC ATGTTCTTC CTGCGTTATC CCCTGATTCT	10260
GTGGATAACC GTATTACCGC CTTGAGTGA GCTGATACCG CTCGCCGCAG CCGAACGACC	10320
GAGCGCAGCG AGTCAGTGAG CGAGGAAGCG GAAGAGCGCC CAATACGCAA ACCGCCTCTC	10380
CCCGCGCGTT GGCGGATTCA TTAATGCAGC TGGCACGACA GGTTCGGCA CTGGAAAGCG	10440
GGCAGTGAGC GCAACGCAAT TAATGTGAGT TAGCTCACTC ATTAGGCACC CCAGGCTTTA	10500
CACTTTATGC TTCCGGCTCG TATGTTGTGT GGAATTGTGA GCGGATAACA ATTTCACACA	10560
GGAAACAGCT ATGACCATGA TTACGCCAAG CTTTGCGAT CAATAATGG ATCACAAACCA	10620
GTATCTCTTA ACGATGTTCT TCGCAGATGA TGATTCAATT TTTAAGTATT TGGCTAGTCA	10680
AGATGATGAA TCTTCATTAT CTGATATATT GCAAATCACT CAATATCTAG ACTTTCTGTT	10740
ATTATTATTG ATCCAATCAA AAAATAAATT AGAAGCCGTG GGTCAATTGTT ATGAATCTCT	10800

TTCAGAGGAA TACAGACAAT TGACAAAATT CACAGACTTT CAAGATTTA AAAAACTGTT	10860
TAACAAGGTC CCTATTGTTA CAGATGGAAG GGTCAAACCT AATAAAGGAT ATTTGTTCGA	10920
CTTGTGATT AGTTTGATGC GATTAAAAAA AGAACCTCT CTAGCTACCA CCGCAATAGA	10980
TCCTGTTAGA TACATAGATC CTCGTCGCAA TATCGCATT TCTAACGTGA TGGATATATT	11040
AAAGTCGAAT AAAGTGAACA ATAATTAATT CTTTATTGTC ATCATGAACG GCGGACATAT	11100
TCAGTTGATA ATCGGCCCA TGTTTCAGG TAAAAGTACA GAATTAATTA GACGAGTTAG	11160
ACGTTATCAA ATAGCTCAAT ATAAATGCGT GACTATAAAA TATTCTAACG ATAATAGATA	11220
CGGAACGGGA CTATGGACGC ATGATAAGAA TAATTTGAA GCATTGGAAG CAACTAAACT	11280
ATGTGATGTC TTGGAATCAA TTACAGATT CTCCGTGATA GGTATCGATG AAGGACAGTT	11340
CTTTCCAGAC ATTGTTGAAT TGATCTCGAT CCCCGAAAT TAATACGACT CACTATAGGG	11400
AGACCACAAAC GGTTTCCCTC TAGCGGGATC AATTCCGCC CTCTCCCTCC CCCCCCCCCTA	11460
ACGTTACTGG CCGAAGCCGC TTGGAATAAG GCCGGTGTGC GTTGTCTAT ATGTTATTTT	11520
CCACCATATT GCCGTCTTT GGCAATGTGA GGGCCCGGAA ACCTGGCCCT GTCTTCTTGA	11580
CGAGCATTCC TAGGGTCTT TCCCCTCTCG CCAAAGGAAT GCAAGGTCTG TTGAATGTCG	11640
TGAAGGAAGC AGTCCTCTG GAAGCTTCTT GAAGACAAAC AACGTCTGTA GCGACCCTTT	11700
GCAGGCAGCG GAACCCCCCA CCTGGCGACA GGTGCCTCTG CGGCCAAAAG CCACGTGTAT	11760
AAGATACACC TGCAAAGCG GCACAACCCC AGTGCCACGT TGTGAGTTGG ATAGTTGTGG	11820
AAAGAGTCAA ATGGCTCTCC TCAAGCGTAT TCAACAAGGG GCTGAAGGAT GCCCAGAAGG	11880
TACCCATTG TATGGGATCT GATCTGGGC CTCGGTGCAC ATGCTTACA TGTGTTAGT	11940
CGAGGTTAAA AAACGTCTAG GCCCCCCGAA CCACGGGGAC GTGGTTTCC TTTGAAAAAC	12000
ACGATAATAC C	12011

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GATCGATGCT AGCCC

15

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GATCGGGCTA GCATC

15

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TTACATGGCC AT

12

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TCACATGGCG AT

12

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TTTGGACTGG GC

12

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TTTGATTGG GC

12

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 14 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TGGTCCTAAT ACTG

14

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 14 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

TGGGCCTAAT ATCG

14

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GCATTATCTA GATGTGTCTT CTGGTCAGAG

30

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CCTGAATTAT AATAATTAAC TGCAGGTCTT

30

(2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 26 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GGGAAAGAAT CCAGAGACAA GAACGG

26

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GGTGAAGTTG TGGATCCATT TGATTG

26

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CAACCTGTAA GGTACCAGCA TCCG

24

(2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GATATGGTGT TAGGCCTTGA TCTGTTTC

27

(2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CGCCATGGAA AAATCAGAGA TCCTCTTCT

29

(2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 31 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

CTGGATCCTA ATTGGAGTTG TTACCCATGT A

31

(2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 23 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

AACCATGGCT GAAAAAGGGA AAA

23

(2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 33 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GGTGAAGCTT AAGATGTGAT TTTACATATT TTA

33

(2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 35 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

AAATAGGATC CCTACAGATC ATTAGATATT AAAAT

35

(2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

CGCCATGGTG TTCAGTGCTT GTTG

24

(2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

CCACAAGCTT AATTAACCAT AATATGCATC A

31

(2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

TTCCCATGGAT TTGGATTGT CTATTGGGT

29

(2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 15462 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

ACCAAAACAAG AGAAGAAA	TGCTTGGTAA TATAAATT	ACTTAAAATT AACTTAGGAT	60
TTAACGACATT GACTAGAAGG	TCAAGAAAAG GGAACTCTAT	AATTCAAAAA ATGTTGAGCC	120
TATTTGATAC ATTTAATGCA	CGTAGGCAAG AAAACATAAC	AAAATCAGCC GGTGGAGCTA	180
TCATTCTGG ACAGAAAAAT	ACTGTCTCTA TATTCGCCCT	TGGACCGACA ATAACGTGATG	240
ATAATGAGAA AATGACATTA	GCTCTTCTAT TTCTATCTCA	TTCACTAGAT AATGAGAAC	300

AACATGCACA	AAGGGCAGGG	TTCTTGGTGT	CTTTATTGTC	AATGGCTTAT	GCCAATCCAG	360
AGCTCTACCT	AACAACAAAT	GGAAGTAATG	CAGATGCCAA	GTATGTCATA	TACATGATTG	420
AGAAAAGATCT	AAAACGGCAA	AAGTATGGAG	GATTGTGGT	TAAGACGAGA	GAGATGATAT	480
ATGAAAAGAC	AACTGATTGG	ATATTGGAA	GTGACCTGGA	TTATGATCAG	GAAACTATGT	540
TGCAGAACGG	CAGGAACAAT	TCAACAATTG	AAGACCTGT	CCACACATTT	GGGTATCCAT	600
CATGTTAGG	AGCTCTTATA	ATACAGATCT	GGATAGTTCT	GGTCAAAGCT	ATCACTAGTA	660
TCTCAGGGTT	AAGAAAAGGC	TTTTTCACCC	GATTGGAAGC	TTTCAGACAA	GATGGAACAG	720
TGCAGGCAGG	GCTGGTATTG	AGCGGTGACA	CAGTGGATCA	GATTGGTCA	ATCATGCGGT	780
CTCAACAGAG	CTTGGTAACT	CTTATGGTTG	AAACATTAAT	AACAATGAAT	ACCAGCAGAA	840
ATGACCTCAC	AACCATAGAA	AAGAATATAC	AAATTGTTGG	CAACTACATA	AGAGATGCAG	900
GTCTCGCTTC	ATTCTTCAT	ACAATCAGAT	ATGGAATTGA	GACCAGAATG	GCAGCTTTGA	960
CTCTATCCAC	TCTCAGACCA	GATATCAATA	GATTAAAAGC	TTTGATGGAA	CTGTATTTAT	1020
CAAAGGGACC	ACCGCGCTCCT	TTCATCTGTA	TCCTCAGAGA	TCCTATACAT	GGTGAGTCG	1080
CACCAGGCAA	CTATCCTGCC	ATATGGAGCT	ATGCAATGGG	GGTGGCAGTT	GTACAAAATA	1140
GAGCCATGCA	ACAGTATGTG	ACGGGAAGAT	CATATCTAGA	CATTGATATG	TTCCAGCTAG	1200
GACAAGCAGT	AGCACGTGAT	GCCGAAGCTC	AAATGAGCTC	AACACTGGAA	GATGAACATTG	1260
GAGTGACACCA	CGAAGCTAAA	GAAAGCTTGA	AGAGACATAT	AAGGAACATA	AACAGTTCAG	1320
AGACATCTT	CCACAAACCG	ACAGGTGGAT	CAGCCATAGA	GATGGCAATA	GATGAAGAGC	1380
CAGAACAAATT	CGAACATAGA	GCAGATCAAG	AACAAAATGG	AGAACCTCAA	TCATCCATAA	1440
TTCAATATGC	CTGGGCAGAA	GGAAATAGAA	GCGATGATCA	GAUTGAGCAA	GCTACAGAAT	1500
CTGACAATAT	CAAGACCGAA	CAACAAAACA	TCAGAGACAG	ACTAAACAAG	AGACTCAACG	1560
ACAAGAAGAA	ACAAAGCAGT	CAACCACCCA	CTAATCCCAC	AAACAGAACA	AACCAGGACG	1620
AAATAGATGA	TCTGTTAAC	GCATTGGAA	GCAACTAAC	GAATCAACAT	TTTAATCTAA	1680
ATCAATAATA	AATAAGAAAA	ACTTAGGATT	AAAGAATCCT	ATCATACCGG	AATATAGGGT	1740
GGTAAATTTA	GAGTCTGCTT	GAAACTCAAT	CAATAGAGAG	TTGATGGAAA	GCGATGCTAA	1800
AAACTATCAA	ATCATGGATT	CTTGGGAAGA	GGAATCAAGA	GATAAAATCAA	CTAATATCTC	1860
CTCGGCCCTC	AACATCATTG	AATTCTACT	CAGCACCGAC	CCCCAAGAAG	ACTTATCGGA	1920
AAACGACACCA	ATCAACACAA	GAACCCAGCA	ACTCAGTGCC	ACCATCTGTC	AACCAGAAAT	1980
CAAACCAACA	GAAACAAGTG	AGAAAGATAG	TGGATCAACT	GACAAAAATA	GACAGTCTGG	2040
GTCATCACAC	GAATGTACAA	CAGAACGAAA	AGATAGAAAC	ATTGATCAGG	AAACTGTACA	2100
GAGAGGACCT	GGGAGAAGAA	GCAGCTCAGA	TAGTAGAGCT	GAGACTGTGG	TCTCTGGAGG	2160
AATCCCCAGA	AGCATCACAG	ATTCTAAAAA	TGGAACCCAA	AACACGGAGG	ATATTGATCT	2220
CAATGAAATT	AGAAAGATGG	ATAAGGACTC	TATTGAGGGG	AAAATGCGAC	AATCTGCAA	2280
TGTTCCAAGC	GAGATATCAG	GAAGTGTGAA	CATATTACAA	ACAGAACAAA	GTAGAACAG	2340
TGATCATGGA	AGAAGCCTGG	AATCTATCAG	TACACCTGAT	ACAAGATCAA	TAAGTGTGTTGT	2400

TACTGCTGCA ACACCAGATG ATGAAGAAGA AATACTAATG AAAAATAGTA GGACAAAGAA	2460
AAGTTCTTCA ACACATCAAG AAGATGACAA AAGAATTAAA AAAGGGGGAA AAGGGAAAGA	2520
CTGGTTAACG AAATCAAAAG ATACCGACAA CCAGATACCA ACATCAGACT ACAGATCCAC	2580
ATCAAAAGGG CAGAAGAAAA TCTCAAAGAC AACAAACCACC AACACCGACA CAAAGGGGCA	2640
AACAGAAATA CAGACAGAAT CATCAGAAAC ACAATCCTCA TCATGGAATC TCATCATCGA	2700
CAACAACACC GACCGGAACG AACAGACAAG CACAACCTCT CCAACAACAA CTTCCAGATC	2760
AACTTATACA AAAGAATCGA TCCGAACAAA CTCTGAATCC AAACCCAAGA CACAAAAGAC	2820
AAATGGAAAG GAAAGGAAGG ATACAGAAGA GAGCAATCGA TTTACAGAGA GGGCAATTAC	2880
TCTATTGCAG AATCTTGGTG TAATTCAATC CACATCAAAA CTAGATTTAT ATCAAGACAA	2940
ACGAGTTGTA TGTGTAGCAA ATGTACTAAA CAATGTAGAT ACTGCATCAA AGATAGATTT	3000
CCTGGCAGGA TTAGTCATAG GGGTTCAAT GGACAACGAC ACAAAATTAA CACAGATACA	3060
AAATGAAATG CTAACACTCA AAGCAGATCT AAAGAAAATG GACGAATCAC ATAGAAGATT	3120
GATAGAAAAT CAAAGAGAAC AACTGTCATT GATCACGTCA CTAATTCAA ATCTCAAAAT	3180
TATGACTGAG AGAGGAGGAA AGAAAGACCA AAATGAATCC AATGAGAGAG TATCCATGAT	3240
CAAAACAAAA TTGAAAGAAG AAAAGATCAA GAAGACCAGG TTTGACCCAC TTATGGAGGC	3300
ACAAGGCATT GACAAGAATA TACCCGATCT ATATCGACAT GCAGGAGATA CACTAGAGAA	3360
CGATGTACAA GTTAAATCAG AGATATTAAG TTCATACAAAT GAGTCAAATG CAACAAGACT	3420
AATACCCAAA AAAGTGAGCA GTACAATGAG ATCACTAGTT GCAGTCATCA ACAACAGCAA	3480
TCTCTCACAA AGCACAAAAC AATCATAACAT AAACGAACCTC AAACGTTGCA AAAATGATGA	3540
AGAAGTATCT GAATTAATGG ACATGTTCAA TGAAGATGTC AACAAATTGCC AATGATCCAA	3600
CAAAGAAACG ACACCGAACAA AACAGACAAG AAACAACAGT AGATCAAAAC CTGTCAACAC	3660
ACACAAAATC AAGCAGAATG AAACAACAGA TATCAATCAA TATACAAATA AGAAAAACTT	3720
AGGATTAAG AATAAATTAA TCCTTGTCCA AAATGAGTAT AACTAACTCT GCAATATACA	3780
CATTCCCAGA ATCATCATTG TCTGAAAATG GTCATATAGA ACCATTACCA CTCAAAGTCA	3840
ATGAACAGAG GAAAGCAGTA CCCCACATTA GAGTTGCCAA GATCGGAAAT CCACCAAAAC	3900
ACGGATCCCG GTATTTAGAT GTCTTCTTAC TCGGCTTCTT CGAGATGGAA CGAATCAAAG	3960
ACAAATACGG GAGTGTGAAT GATCTCGACA GTGACCCGAG TTACAAAGTT TGTGGCTCTG	4020
GATCATTACC AATCGGATTG GCTAAGTACA CTGGGAATGA CCAGGAATTG TTACAAGCCG	4080
CAACCAAACT GGATATAGAA GTGAGAAGAA CAGTCAAAGC GAAAGAGATG GTTGTTCACA	4140
CGGTACAAAA TATAAAACCA GAACTGTACC CATGGTCAA TAGACTAAGA AAAGGAATGC	4200
TGTTCGATGC CAACAAAGTT GCTCTTGCTC CTCATGTCT TCCACTAGAT AGGAGCATAA	4260
AATTTAGAGT AATCTTCGTG AATTGTACGG CAATTGGATC AATAACCTTG TTCAAAATTC	4320
CTAAGTCAAT GGCATCACTA TCGTTAACCA ACACAATATC AATCAATCTG CAGGTACACA	4380
AAAAAACAGG GGTTCAGACT GATTCTAAAG GGATAGTTCA AATTTGGAT GAGAAAGGCG	4440
AAAAATCACT GAATTCATG GTCCATCTCG GATTGATCAA AAGAAAAGTA GGCAGAATGT	4500

ACTCTGTTGA ATACTGTAAA CAGAAAATCG AGAAAATGAG ATTGATATTT TCTTTAGGAC	4560
TAGTTGGAGG AATCAGTCTT CATGTCAATG CAACTGGTC CATATCAAAA ACACTAGCAA	4620
GTCAGCTGGT ATTCAAAAGA GAGATTGTT ATCCTTTAAT GGATCTAAAT CCGCATCTCA	4680
ATCTAGTTAT CTGGGCTTCA TCAGTAGAGA TTACAAGAGT GGATGCAATT TTCCAACCTT	4740
CTTTACCTGG CGAGTTCAGA TACTATCCTA ATATTATTGC AAAAGGAGTT GGGAAAATCA	4800
AACAATGGAA CTAGTAATCT CTATTTAGT CCGGACGTAT CTATTAAGCC GAAGCAAATA	4860
AAGGATAATC AAAAACCTAG GACAAAAGAG GTCAATACCA ACAACTATTA GCAGTCACAC	4920
TCGCAAGAAT AAGAGAGAAG GGACAAAAAA AGTCAAATAG GAGAAATCAA AACAAAAGGT	4980
ACAGAACACC AGAACAAACAA AATCAAAACA TCCAECTCAC TCAAAACAAA AATTCCAAA	5040
GAGACCGGCA ACACAACAAG CACTGAACAC AATGCCAATC TCAATACTGC TAATTATTAC	5100
AACCATGATC ATGGCATCTT TCTGCCAAT AGATATCACA AAACATACAGC ACGTAGGTGT	5160
ATTGGTCAAC AGTCCCCAAAG GGATGAAGAT ATCACAAAAC TTTGAAACAA GATATCTAAT	5220
TTTGAGCCTC ATACCAAAAA TAGAAGACTC TAACTCTTGT GGTGACCAAC AGATCAAGCA	5280
ATACAAGAAG TTATTGGATA GACTGATCAT CCCTTTATAT GATGGATTAA GATTACAGAA	5340
AGATGTGATA GTAACCAATC AAGAATCCAA TGAAAACACT GATCCCAGAA CAAAACGATT	5400
CTTTGGAGGG GTAATTGGAA CCATTGCTCT GGGAGTAGCA ACCTCAGCAC AAATTACAGC	5460
GGCAGTTGCT CTGGTTGAAG CCAAGCAGGC AAGATCAGAC ATCGAAAAC TCAAAGAACG	5520
AATTAGGGAC ACAAACAAAG CAGTGCAGTC AGTCAGAGC TCCATAGGAA ATTTAATAGT	5580
AGCAATTAAA TCAGTCCAGG ATTATGTTAA CAAAGAAATC GTGCCATCGA TTGCGAGGCT	5640
AGGTTGTGAA GCAGCAGGAC TTCAATTAGG AATTGCATTA ACACAGCATT ACTCAGAATT	5700
AAACAAACATA TTTGGTGATA ACATAGGATC GTTACAAGAA AAAGGAATAA AATTACAAGG	5760
TATAGCATCA TTATACCGCA CAAATATCAC AGAAATATTC ACAACATCAA CAGTTGATAA	5820
ATATGATATC TATGATCTGT TATTACAGA ATCAATAAAG GTGAGAGTTA TAGATGTTGA	5880
CTTGAATGAT TACTCAATCA CCCTCCAAGT CAGACTCCCT TTATTAACCA GGCTGCTGAA	5940
CACTCAGATC TACAAAGTAG ATTCCATATC ATATAACATC CAAAACAGAG AATGGTATAT	6000
CCCTCTCCCC AGCCATATCA TGACGAAAGG GGCATTTCTA GGTGGAGCAG ACGTCAAAGA	6060
ATGTATAGAA GCATTCAGCA GCTATATATG CCCTTCTGAT CCAGGATTTG TATTAAACCA	6120
TGAAATAGAG AGCTGTTAT CAGGAAACAT ATCCCAATGT CCAAGAACAA CGGTACACATC	6180
AGACATTGTT CCAAGATATG CATTGTCAA TGGAGGAGTG GTTGCAAACACT GTATAACAAC	6240
CACCTGTACA TGCAACGGAA TTGGTAATAG AATCAATCAA CCACCTGATC AAGGAGTAAA	6300
AATTATAACA CATAAAGAAT GTAGTACTGT GGGTATCAAC GGAATGCTGT TCAATACAAA	6360
TAAAGAAGGA ACTCTTGCAT TCTATACACC AAATGATATA ACACAAACAA ATTCTGTTAC	6420
ACTGGATCCA ATTGACATAT CAATCGAGCT CAACAAGGCC AAATCAGATC TAGAAGAAC	6480
AAAAGAATGG ATAAGAAGGT CAAATCAAAA ACTAGATTCT ATTGGAAATT GGCATCAATC	6540
TAGCACTACA ATCATAATTA TTTGATAAT GATCATTATA TTGTTTATAA TTAATATAAC	6600

GATAATTACA	ATGGCAATTAA	AGTATTACAG	AATTCAAAG	AGAAATCGAG	TGGATCAAAA	6660
TGACAAGCCA	TATGTACTAA	CAAACAAATA	ACATATCTAC	AGATCATTAG	ATATTAAAAT	6720
TATAAAAAAC	TTAGGAGTAA	AGTTACGCAA	TCCAECTCTA	CTCATATAAT	TGAGGAAGGA	6780
CCCAATAGAC	AAATCCAAAT	TCGAGATGGA	ATACTGGAAG	CATACCAATC	ACGGAAAGGA	6840
TGCTGGTAAT	GAGCTGGAGA	CGTCTATGGC	TACTCATGGC	AACAAGCTCA	CTAATAAGAT	6900
AATATACATA	TTATGGACAA	TAATCCTGGT	GTTATTATCA	ATAGTCTTCA	TCATAGTGCT	6960
AATTAATTCC	ATCAAAAGTG	AAAAGGCCA	CGAACATCATTG	CTGCAAGACA	TAAATAATGA	7020
GTTTATGGAA	ATTACAGAAA	AGATCCAAAT	GGCATCGGAT	AATACCAATG	ATCTAATACAA	7080
GTCAGGAGTG	AATACAAGGC	TTCTTACAAT	TCAGAGTCAT	GTCCAGAATT	ACATACCAAT	7140
ATCATTGACA	CAACAGATGT	CAGATCTTAG	GAAATTCAATT	AGTGAAATTAA	CAATTAGAAA	7200
TGATAATCAA	GAAGTGCTGC	CACAAAGAAT	AACACATGAT	GTAGGTATAA	AACCTTTAAA	7260
TCCAGATGAT	TTTGGAGAT	GCACGCTCTGG	TCTTCCATCT	TTAATGAAAA	CTCCAAAAAT	7320
AAGGTTAATG	CCAGGGCCGG	GATTATTAGC	TATGCCAACG	ACTGTTGATG	GCTGTGTTAG	7380
AACTCCGTCT	TTAGTTATAA	ATGATCTGAT	TTATGCTTAT	ACCTCAAATC	TAATTACTCG	7440
AGGTTGTCAG	GATATAGGAA	AATCATATCA	AGTCTTACAG	ATAGGGATAA	TAACTGTAAA	7500
CTCAGACTTG	GTACCTGACT	TAAATCCTAG	GATCTCTCAT	ACCTTTAACAA	TAAATGACAA	7560
TAGGAAGTCA	TGTTCTCTAG	CACTCCTAAA	TATAGATGTA	TATCAACTGT	GTTCAACTCC	7620
CAAAGTTGAT	GAAAGATCAG	ATTATGCATC	ATCAGGCATA	GAAGATATTG	TACTTGATAT	7680
TGTCAATTAT	GATGGTTCAA	TCTCAACAAAC	AAGATTTAAG	AATAATAACA	TAAGCTTTGA	7740
TCAACCATAT	GCTGCACTAT	ACCCATCTGT	TGGACCAGGG	ATATACTACA	AAGGCAAAAT	7800
AATATTCTC	GGGTATGGAG	GTCTGAAACA	TCCAATAAT	GAGAATGTAA	TCTGCAACAC	7860
AACTGGGTGC	CCCGGGAAAA	CACAGAGAGA	CTGTAATCAA	GCATCTCATA	GTACTTGGTT	7920
TTCAGATAGG	AGGATGGTCA	ACTCCATCAT	TGTGGCTGAC	AAAGGCTTAA	ACTCAATTCC	7980
AAAATTGAAA	GTATGGACGA	TATCTATGCG	ACAAAATTAC	TGGGGTCAG	AAGGAAGGTT	8040
ACTTCTACTA	GGTAACAAGA	TCTATATATA	TACAAGATCT	ACAAGTTGGC	ATAGCAAGTT	8100
ACAATTAGGA	ATAATTGATA	TTACTGATTA	CAGTGATATA	AGGATAAAAT	GGACATGGCA	8160
TAATGTGCTA	TCAAGACCAG	GAAACAATGA	ATGTCCATGG	GGACATTCTAT	GTCCAGATGG	8220
ATGTATAACA	GGAGTATATA	CTGATGCATA	TCCACTCAAT	CCCACAGGG	GCATTGTGTC	8280
ATCTGTCATA	TTAGACTCAC	AAAAATCGAG	AGTGAACCCA	GTCATAACTT	ACTCAACAGC	8340
AACCGAAAGA	GTAAACGAGC	TGGCCATCCT	AAACAGAACAA	CTCTCAGCTG	GATATACAAC	8400
AACAAGCTGC	ATTACACACT	ATAACAAAGG	ATATTGTTTT	CATATAGTAG	AAATAAAATCA	8460
TAAAAGCTTA	AACACATTTC	AACCCATGTT	GTTCAAAACA	GAGATTCCAA	AAAGCTGCAG	8520
TTAACATCAA	TTAACCATCAA	TCTATCTATA	ATACAAGTAT	ATGATAAGTA	ATCAGCAATC	8580
AGACAATAGA	CAAAAGGGAA	ATATAAAAAAA	CTTAGGAGCA	AAGCGTGCTC	8640	
GGGAAATGGA	CACTGAATCT	AACAATGGCA	CTGTATCTGA	CATACTCTAT	CCTGAGTGTC	8700

ACCTTAACTC TCCTATCGTT AAAGGTAAAA TAGCACAATT ACACACTATT ATGAGTCTAC	8760
CTCAGCCTTA TGATATGGAT GACGACTCAA TACTAGTTAT CACTAGACAG AAAATAAAAC	8820
TTAATAAATT GGATAAAAAGA CAACGATCTA TTAGAAGATT AAAATTAATA TTAACTGAAA	8880
AAGTGAATGA CTTAGGAAAA TACACATTAA TCAGATATCC AGAAATGTCA AAAGAAATGT	8940
TCAAATTATA TATACCTGGT ATTAACAGTA AAGTGACTGA ATTATTACTT AAAGCAGATA	9000
GAACATATAG TCAAATGACT GATGGATTAA GAGATCTATG GATTAATGTG CTATCAAAAT	9060
TAGCCTCAAA AAATGATGGA AGCAATTATG ATCTTAATGA AGAAATTAAT AATATATCGA	9120
AAGTCACAC AACCTATAAA TCAGATAAT GGTATAATCC ATTCAAAACA TGGTTTACTA	9180
TCAAGTATGA TATGAGAAGA TTACAAAAAG CTCGAAATGA GATCACTTT AATGTTGGGA	9240
AGGATTATAA CTTGTTAGAA GACCAGAAGA ATTTCTTATT GATACATCCA GAATTGGTTT	9300
TGATATTAGA TAAACAAAAC TATAATGGTT ATCTAATTAC TCCTGAATTA GTATTGATGT	9360
ATTGTGACGT AGTCGAAGGC CGATGGAATA TAAGTGCATG TGCTAAGTTA GATCCAAAT	9420
TACAATCTAT GTATCAGAAA GGTAATAACC TGTGGGAAGT GATAGATAAA TTGTTTCCAA	9480
TTATGGGAGA AAAGACATTG GATGTGATAT CGTTATTAGA ACCACTTGCA TTATCCTTAA	9540
TTCAAACCTCA TGATCCTGTT AAACAACAA GAGGAGCTTT TTTAAATCAT GTGTTATCCG	9600
AGATGGAATT AATATTTGAA TCTAGAGAAT CGATTAAGGA ATTTCTGAGT GTAGATTACA	9660
TTGATAAAAT TTTAGATATA TTTAATAAGT CTACAATAGA TGAAATAGCA GAGATTTCT	9720
CTTTTTTAG AACATTTGGG CATCCTCCAT TAGAAGCTAG TATTGCAGCA GAAAAGGTTA	9780
GAAAATATAT GTATATTGGA AAACAATTAA AATTGACAC TATTAATAAA TGTCATGCTA	9840
TCTTCTGTAC AATAATAATT AACGGATATA GAGAGAGGCA TGGTGGACAG TGGCCTCCTG	9900
TGACATTACC TGATCATGCA CACGAATTCA TCATAAATGC TTACGGTTCA AACTCTGCGA	9960
TATCATATGA AAATGCTGTT GATTATTACC AGAGCTTTAT AGGAATAAAA TTCAATAAAT	10020
TCATAGAGCC TCAGTTAGAT GAGGATTGA CAATTTATAT GAAAGATAAA GCATTATCTC	10080
CAAAAAAATC AAATTGGGAC ACAGTTTATC CTGCATCTAA TTTACTGTAC CGTACTAACG	10140
CATCCAACGA ATCACGAAGA TTAGTTGAAG TATTTATAGC AGATAGTAAA TTTGATCCTC	10200
ATCAGATATT GGATTATGTA GAATCTGGG ACTGGTTAGA TGATCCAGAA TTTAATATTT	10260
CTTATAGTCT TAAAGAAAAA GAGATCAAAC AGGAAGGTAG ACTCTTGCA AAAATGACAT	10320
ACAAAATGAG AGCTACACAA GTTTTATCAG AGACCTACT TGCAAATAAC ATAGGAAAAT	10380
TCTTTCAAGA AAATGGGATG GTGAAGGGAG AGATTGAATT ACTTAAGAGA TTAACAACCA	10440
TATCAATATC AGGAGTTCCA CGGTATAATG AAGTGTACAA TAATTCTAAA AGCCATACAG	10500
ATGACCTTAA AACCTACAAT AAAATAAGTA ATCTTAATTT GTCTTCTAAT CAGAAATCAA	10560
AGAAATTTGA ATTCAAGTCA ACGGATATCT ACAATGATGG ATACGAGACT GTGAGCTGTT	10620
TCCTAACAAAC AGATCTCAA AAATACTGTC TTAATTGGAG ATATGAATCA ACAGCTCTAT	10680
TTGGAGAAAC TTGCAACCAA ATATTGGAT TAAATAAATT GTTTAATTGG TTACACCCTC	10740
GTCTTGAAGG AAGTACAATC TATGTAGGTG ATCCTTACTG TCCTCCATCA GATAAAGAAC	10800

ATATATCATT AGAGGATCAC CCTGATTCTG GTTTTACGT TCATAACCCA AGAGGGGTA	10860
TAGAAGGATT TTGTCAAAAA TTATGGACAC TCATATCTAT AAGTGAATA CATCTAGCAG	10920
CTGTTAGAAC AGGCCTGAGG GTGACTGCAA TGGTTCAAGG AGACAATCAA GCTATAGCTG	10980
TAACCACAAG AGTACCCAAC AATTATGACT ACAGAGTTAA GAAGGAGATA GTTTATAAAG	11040
ATGTAGTGAG ATTTTTGAT TCATTAAGAG AAGTGTGGA TGATCTAGGT CATGAACCTA	11100
AATTAAATGA AACGATTATA AGTAGCAAGA TGTTCATATA TAGCAAAAGA ATCTATTATG	11160
ATGGGAGAAT TCTTCCTCAA GCTCTAAAAG CATTATCTAG ATGTGTCTTC TGGTCAGAGA	11220
CAGTAATAGA CGAAACAAGA TCAGCATCTT CAAATTTGGC AACATCATTT GCAGAAAGCAA	11280
TTGAGAATGG TTATTCACCT GTTCTAGGAT ATGCATGCTC AATTTTAAG AATATTCAAC	11340
AACTATATAT TGCCCTTGGG ATGAATATCA ATCCAACAT AACACAGAAT ATCAGAGATC	11400
AGTATTTAG GAATCCAAAT TGGATGCAAT ATGCCTCTT AATACCTGCT AGTGTGAGGG	11460
GATTCAATCA CATGGCGATG TCAAGATGTT TTGTAAGGAA TATTGGTGAT CCATCAGTTG	11520
CCGCATTGGC TGATATTAAA AGATTATTA AGGCGAATCT ATTAGACCGA AGTGTGAGGG	11580
ATAGGATTAT GAATCAAGAA CCAGGTGAGT CATCTTTTG TGATTGGGCT TCAGATCCAT	11640
ATTCATGCAA TTTACCACAA TCTCAAAATA TAACCACCAT GATAAAAAT ATAACAGCAA	11700
GGAATGTATT ACAAGATTCA CCAAATCCAT TATTATCTGG ATTATTCACA AATACAATGA	11760
TAGAAGAAGA TGAAGAATTAA GCTGAGTTCC TGATGGACAG GAAGGTAATT CTCCCTAGAG	11820
TTGCACATGA TATTCTAGAT AATTCTCTCA CAGGAATTAG AAATGCCATA GCTGGAAATGT	11880
TAGATACGAC AAAATCACTA ATTGGGTTG GCATAAATAG AGGAGGACTG ACATATAGTT	11940
TGTTGAGGAA AATCAGTAAT TACGATCTAG TACAATATGA AACACTAAGT AGGACTTTGC	12000
GACTAATTGT AAGTGATAAA ATCAAGTATG AAGATATGTG TTCGGTAGAC CTTGCCATAG	12060
CATTGCGACA AAAGATGTGG ATTCAATTAT CAGGAGGAAG GATGATAAGT GGACTTGAAA	12120
CGCCTGACCC ATTAGAATTAA CTATCTGGG TAGTAATAAC AGGATCAGAA CATTGTAAAA	12180
TATGTTATTC TTCAGATGGC ACAAAACCAT ATACTTGGAT GTATTTACCC GGTAATATCA	12240
AAATAGGATC AGCAGAAACA GGTATATCGT CATTAAGAGT TCCTTATTTT GGATCAGTCA	12300
CTGATGAAAG ATCTGAAGCA CAATTAGGAT ATATCAAGAA TCTTAGTAAA CCTGCAAAG	12360
CCGCAATAAG AATAGCAATG ATATATACAT GGGCATTGG TAATGATGAG ATATCTTGG	12420
TGGAAGCCTC ACAGATAGCA CAAACACGTG CAAATTTAC ACTAGATAGT CTCAAAATTT	12480
TAACACCGGT AGCTACATCA ACAAAATTAT CACACAGATT AAAGGATACT GCAACTCAGA	12540
TGAAATTCTC CAGTACATCA TTGATCAGAG TCAGCAGATT CATAACAATG TCCAATGATA	12600
ACATGTCTAT CAAAGAAGCT AATGAAACCA AAGATACTAA TCTTATTTT CAACAAATAA	12660
TGTTAACAGG ATTAAGTGTG TTCGAATATT TATTTAGATT AAAAGAAACC ACAGGACACA	12720
ACCCTATAGT TATGCATCTG CACATAGAAG ATGAGTGTG TATTAAGAA AGTTTTAATG	12780
ATGAACATAT TAATCCAGAG TCTACATTAG AATTAATTG ATATCCTGAA AGTAATGAAT	12840
TTATTTATGA TAAAGACCCA CTCAAAGATG TGGACTTATC AAAACTTATG GTTATTAAG	12900

ACCATTCTTA CACAATTGAT ATGAATTATT GGGATGATAC TGACATCATA CATGCAATT	12960
CAATATGTAC TGCAATTACA ATAGCAGATA CTATGTCACA ATTAGATCGA GATAATTAA	13020
AAGAGATAAT AGTTATTGCA AATGATGATG ATATTAATAG CTTAACACT GAATTTTGA	13080
CTCTTGACAT ACTTGTATTT CTCAAGACAT TTGGTGGATT ATTAGTAAAT CAATTGCAT	13140
ACACTCTTTA TAGTCTAAA ATAGAAGGTA GGGATCTCAT TTGGGATTAT ATAATGAGAA	13200
CACTGAGAGA TACTTCCCAT TCAATATTAA AAGTATTATC TAATGCATTA TCTCATCCTA	13260
AAGTATTCAA GAGGTTCTGG GATTGTGGAG TTTTAAACCC TATTATGGG CCTAATATCG	13320
CTAGTCAAGA CCAGATAAAA CTTGCCCTAT CTATATGTGA ATATTCACTA GATCTATTTA	13380
TGAGAGAATG GTTGAATGGT GTATCACTTG AAATATACAT TTGTGACAGC GATATGGAAG	13440
TTGCAAATGA TAGGAAACAA GCCTTATTT CTAGACACCT TTCATTGTT TGTTGTTAG	13500
CAGAAATTGC ATCTTCGGA CCTAACCTGT TAAACTTAAC ATACTTGGAG AGACTTGATC	13560
TATTGAAACA ATATCTTGA TTAAATATTA AAGAAGACCC TACTCTAAA TATGTACAAA	13620
TATCTGGATT ATTAATTAAA TCGTCCCCT AACTGTAAC ATACGTAAGA AAGACTGCAA	13680
TCAAATATCT AAGGATTCGC GGTATTAGTC CACCTGAGGT AATTGATGAT TGGGATCCGG	13740
TAGAAGATGA AAATATGCTG GATAACATTG TCAAAACTAT AAATGATAAC TGTAATAAAG	13800
ATAATAAAGG GAATAAAATT ACAATTCT GGGGACTAGC ACTTAAGAAC TATCAAGTCC	13860
TTAAAATCAG ATCTATAACA AGTGATTCTG ATGATAATGA TAGACTAGAT GCTAATACAA	13920
GTGGTTGAC ACTTCCTCAA GGAGGGAATT ATCTATCGCA TCAATTGAGA TTATTGGAA	13980
TCAACAGCAC TAGTTGTCTG AAAGCTCTG AGTTATCACA AATTTAATG AAGGAAGTCA	14040
ATAAAGACAA GGACAGGCTC TTCCTGGAG AAGGAGCAGG AGCTATGCTA GCATGTTATG	14100
ATGCCACATT AGGACCTGCA GTTAATTATT ATAATTCAAG TTTGAATATA ACAGATGTAA	14160
TTGGTCAACG AGAATTGAAA ATATTCCTT CAGAGGTATC ATTAGTAGGT AAAAATTAG	14220
GAAATGTGAC ACAGATTCTT AACAGGGTAA AAGTACTGTT CAATGGGAAT CCTAATTCAA	14280
CATGGATAGG AAATATGGAA TGTGAGAGCT TAATATGGAG TGAATTAAAT GATAAGTCCA	14340
TTGGATTAGT ACATTGTGAT ATGGAAGGAG CTATCGTAA ATCAGAAGAA ACTGTTCTAC	14400
ATGAACATTA TAGTGTATA AGAATTACAT ACTTGATTGG GGATGATGAT GTTGTAG	14460
TTTCCAAAT TATACCTACA ATCACTCCGA ATTGGTCTAG AATACTTTAT CTATATAAAT	14520
TATATTGGAA AGATGTAAGT ATAATATCAC TCAAAACTTC TAATCCTGCA TCAACAGAAT	14580
TATATCTAAT TTCGAAAGAT GCATATTGTA CTATAATGGA ACCTAGTGAA ATTGTTTAT	14640
CAAAACTTAA AAGATTGTCA CTCTTGGAG AAAATAATCT ATTAATGGG ATCATTAT	14700
CAAAGAAGAG GAATAATGAA TGGTACATC ATGAAATCAA AGAAGGAGAA AGAGATTATG	14760
GAATCATGAG ACCATATCAT ATGGCACTAC AAATCTTGG ATTCAAATC AATTTAAATC	14820
ATCTGGCGAA AGAATTGTTA TCAACCCAG ATCTGACTAA TATCAACAAAT ATAATCCAAA	14880
GTTCAGCG AACAAATAAG GATGTTTAT TTGAATGGAT TAATATAACT CATGATGATA	14940
AGAGACATAA ATTAGGCGGA AGATATAACA TATTCCACT GAAAAATAAG GGAAAGTTAA	15000

GACTGCTATC GAGAAGACTA GTATTAAGTT GGATTCATT ATCATTATCG ACTCGATTAC	15060
TTACAGGTCG CTTTCCTGAT GAAAAATTTG AACATAGAGC ACAGACTGGA TATGTATCAT	15120
TAGCTGATAC TGATTTAGAA TCATTAAGT TATTGTCGAA AACATCATT AAGAATTACA	15180
GAGAGTGTAT AGGATCAATA TCATATTGGT TTCTAACCAA AGAAGTTAAA ATACTTATGA	15240
AATTGATCGG TGGTGCTAAA TTATTAGGAA TTCCCAGACA ATATAAAGAA CCCGAAGACC	15300
AGTTATTAGA AAACATACAAT CAACATGATG AATTGATAT CGATTAAAAC ATAAATACAA	15360
TGAAGATATA TCCTAACCTT TATCTTAAG CCTAGGAATA GACAAAAAGT AAGAAAAACA	15420
TGTAATATAT ATATACCAAA CAGAGTTCTT CTCTGTTG GT	15462

(2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TTGTCTGGGA AT

12

(2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

TTGCCTGGGA AT

12

(2) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

TTGTTTGGGA AT

12

(2) INFORMATION FOR SEQ ID NO:49:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

TTGTCTGGTA AT

12

(2) INFORMATION FOR SEQ ID NO:50:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

AACTTTAAAT TA

12

(2) INFORMATION FOR SEQ ID NO:51:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

AACTTAAAAT TA

12

(2) INFORMATION FOR SEQ ID NO:52:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

TTAAAGACAT TG

12

(2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

TTTAAGACAT TG

12

(2) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GCAGATGTCA AG

12

(2) INFORMATION FOR SEQ ID NO:55:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

GCAGATGCCA AG

12

(2) INFORMATION FOR SEQ ID NO:56:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

CGAATCTAAA GA

12

(2) INFORMATION FOR SEQ ID NO:57:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

CGAAGCTAAA GA

12

(2) INFORMATION FOR SEQ ID NO:58:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

GAAATATTGA TC

12

(2) INFORMATION FOR SEQ ID NO:59:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

GAAACATTGA TC

12

(2) INFORMATION FOR SEQ ID NO:60:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

TCTCTACCCA AC

12

(2) INFORMATION FOR SEQ ID NO:61:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

TCGTTAACCA AC

12

(2) INFORMATION FOR SEQ ID NO:62:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

AGTACAATAG GT

12

(2) INFORMATION FOR SEQ ID NO:63:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

AGTACTGTGG GT

12

(2) INFORMATION FOR SEQ ID NO:64:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

GCACTTGATC CA

12

(2) INFORMATION FOR SEQ ID NO:65:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

ACACTGGATC CA

12

(2) INFORMATION FOR SEQ ID NO:66:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

CCATCATTGT TGTTGACAA

19

(2) INFORMATION FOR SEQ ID NO:67:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

CCATCATTGT GGCTGACAA

19

(2) INFORMATION FOR SEQ ID NO:68:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 11 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

TTACATGGCC A

11

(2) INFORMATION FOR SEQ ID NO:69:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 11 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

TCACATGGCG A

11

(2) INFORMATION FOR SEQ ID NO:70:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 12 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

TTTGGACTGG GC

12

(2) INFORMATION FOR SEQ ID NO:71:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

TTTGATTGG GC

12

(2) INFORMATION FOR SEQ ID NO:72:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

GGTCCTAATA CT

12

(2) INFORMATION FOR SEQ ID NO:73:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

GGGCCTAATA TC

12

(2) INFORMATION FOR SEQ ID NO:74:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74

CCATAGAGAG TCCATGGAAA GCGACGCTAA AAACTATC

38